COMPETITIVE ASSESSMENT OF THE ENERGY MARKET IN NEW ENGLAND

David B. Patton, Ph.D.
Robert A. Sinclair, Ph.D.
Pallas M. LeeVanSchaick

POTOMAC ECONOMICS, LTD.

Independent Market Advisor
to ISO New England

May 2002

TABLE OF CONTENTS

I. Executive Summary

This report evaluates the competitive performance of the New England wholesale electricity market during 2001 by examining whether conduct of market participants was consistent with workable competition. In particular, the report seeks to identify attempts to exercise market power by withholding generating resources from the market. This identification is subject to some uncertainty because observed conduct that is consistent with an attempt to exercise market power is, in many cases, compatible with competitive behavior. For example, legitimate forced outages of generating resources occur even in the most competitive markets.

Therefore, we employ an empirical analysis that is intended to differentiate anticompetitive withholding of resources from conduct that is competitively justified. This analysis consistently indicates that the New England markets have been workably competitive and produces little evidence of persistent economic or physical withholding. While the analysis shows that the New England market has not been subject to systematic withholding, it cannot exclude the possibility that discrete instances of physical withholding occurred via specific outages or deratings.

Therefore, ISO New England's physical audit program, designed to verify that outages and significant deratings are legitimate, remains an important program to detect and deter this conduct. Likewise, monitoring and mitigation of economic withholding continues to be an important function to ensure that if and when a supplier has market power, that attempts to exercise it will be detected and effectively addressed.

Other findings of the withholding analysis conducted for this report relate to out-of-merit dispatch and the implementation of three-part bidding. First, the report finds that the apparent economic withholding by generators that are frequently dispatched out-of-merit is consistent with the incentives provided by the out-of-merit pricing rules. In some cases, this apparent withholding may reflect locational market power that may exist when market areas with few competitors are isolated by transmission constraints. The ISO screens for and mitigates this form of market power under Market Rule 17. Differentiating between these two factors may not be possible and is not within the scope of this report. Hence, the market power findings do not extend to locational market power.

However, the findings regarding out-of-merit bidding incentives do serve to emphasize one of the benefits of the locational clearing price system to be implemented as part of New England's standard market design. Under SMD, suppliers in constrained areas will have the incentive, absent market power, to bid their marginal costs. This should improve the efficiency of the dispatch, as well as establish more accurate economic price signals in the constrained areas.

With regard to three-part bidding, the report finds that energy bids more closely reflected competitive levels (i.e., the amount of potential economic withholding detected was reduced) after the implementation of three-part bidding, which allows suppliers to separately bid their start-up and no-load costs. Although other factors may have changed over the same time-frame, this result is consistent with the expectation that providing generators this additional flexibility has allowed them more accurately to represent their marginal costs in the bids.

In addition to the withholding analysis, the report analyzes the highest-priced hours during the summer of 2001 to determine the extent to which inefficient market rules or procedures, unjustified actions by the ISO, or withholding by participants may have contributed to inflated price levels in these hours. Based on this analysis the report finds that:

- In the majority of the high-priced hours, the prices were warranted based on the deficiency in internal resources that existed in New England, and the ISO's actions in response to these deficiencies were consistent with the market rules and procedures.
- The market rules that prevailed during 2001 may have set prices at unjustifiably high levels during some periods when a deficiency did not exist. This concern has been addressed by the pricing reforms recently implemented by the ISO.
- The current market rules imply that all reserves are infinitely valuable, which contributes to setting extremely high energy prices when costly actions are taken to maintain the reserves. This would most appropriately be addressed by assigning an explicit value to various types and quantities of reserves (i.e., establishing a demand function for reserves).
- The New York ISO market rules related to scheduled exports from New York to New England did restrict New England's access to lower cost imports in a few hours. The New York ISO is implementing changes to its market models that should minimize this possibility in the future.
- No clear evidence was found that economic or physical withholding substantially contributed to inflating the energy prices in these hours. However, discrete instances of physical withholding could not be excluded, which supports the need to continue the monitoring audits of outages and substantial deratings.

Identifying and Evaluating Strategic Withholding

Market power in electricity markets is generally exercised by withholding supplies in an attempt to raise the market-clearing price. A resource can be physically withheld by claiming the unit is unavailable or not offering it when it would be economic to run, or economically withheld by bidding the unit at a price higher than its marginal cost (including opportunity costs) to reduce the unit's output.

The critical task in a withholding analysis is to differentiate strategic withholding from competitive conduct that could appear to be physical or economic withholding. For example, a forced outage of a generating unit may be either legitimate or a strategic attempt to raise prices by physically withholding the unit. To differentiate between these two alternatives, this report evaluates the potential withholding in light of the market conditions and participant characteristics that would tend to create the ability and incentive to exercise market power.

Based on the economic theory described in the report, the two key factors that should be correlated to market power are 1) market participant size and 2) high demand conditions. The intuition regarding the size of the participant is straightforward. Large suppliers have not only more resources to withhold, but also a larger quantity of other resources that would benefit from the higher energy prices. The importance of the high demand conditions relates to the nature of the market supply (i.e., the supply curve).

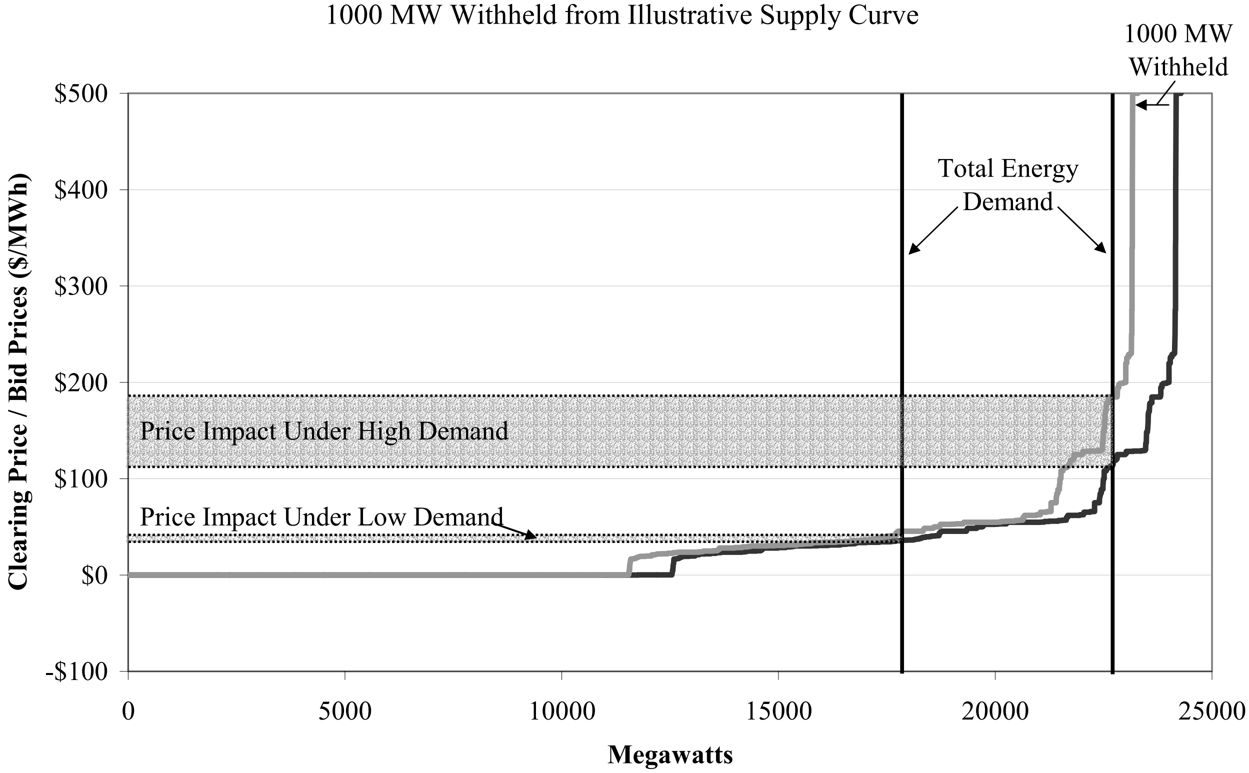
The supply curve that characterizes the energy bid costs of the resources in the market tends to be relatively flat at low output levels, but becomes very steep at times when only the most expensive units are available to meet incremental demand. This characteristic is shown below in Figure E-1, which is a supply curve for the New England electricity market. The shape of the supply curve is important for two reasons.

First, even in the absence of market power, markets that suffer from capacity shortages in peak periods may experience considerable price fluctuations. Since electricity cannot be economically stored in large quantities, higher-cost resources must be used to meet demand at peak hours. Therefore, one cannot draw conclusions regarding the competitive performance of the market from price fluctuations alone.

Second, prices will be much more sensitive to shifts in the available supply when the market is clearing in the relatively steep portion of the supply curve. Therefore, strategic withholding intended to raise prices will generally have larger effects under these conditions. Likewise, these conditions provide the greatest incentive for a supplier with market power to withhold. Fortunately, the vast majority of hours exhibit load levels that correspond to the relatively flat (i.e., more elastic) portion of the supply curve where there is very little incentive to withhold.

Figure E-1 illustrates these characteristics by showing how the price effects of withholding 1000 MW vary at different points on the supply curve. This figure uses actual bids and output for the New England market on August 9, 2001.

Figure E-1
Impacts of Withholding Under High and Low Demand Conditions
1000 MW Withheld from Illustrative Supply Curve



This report does not address locational market power that is associated with transmission congestion for two reasons. First, the New England ISO manages congestion by dispatching units out-of-merit and their bids are screened for potential market power and mitigated when the screens are failed. Second, the out-of-merit congestion management approach compromises suppliers' incentive to bid their marginal costs, making distinguishing attempts to exercise market power from competitive bids very difficult.

Summary of the Withholding Analysis

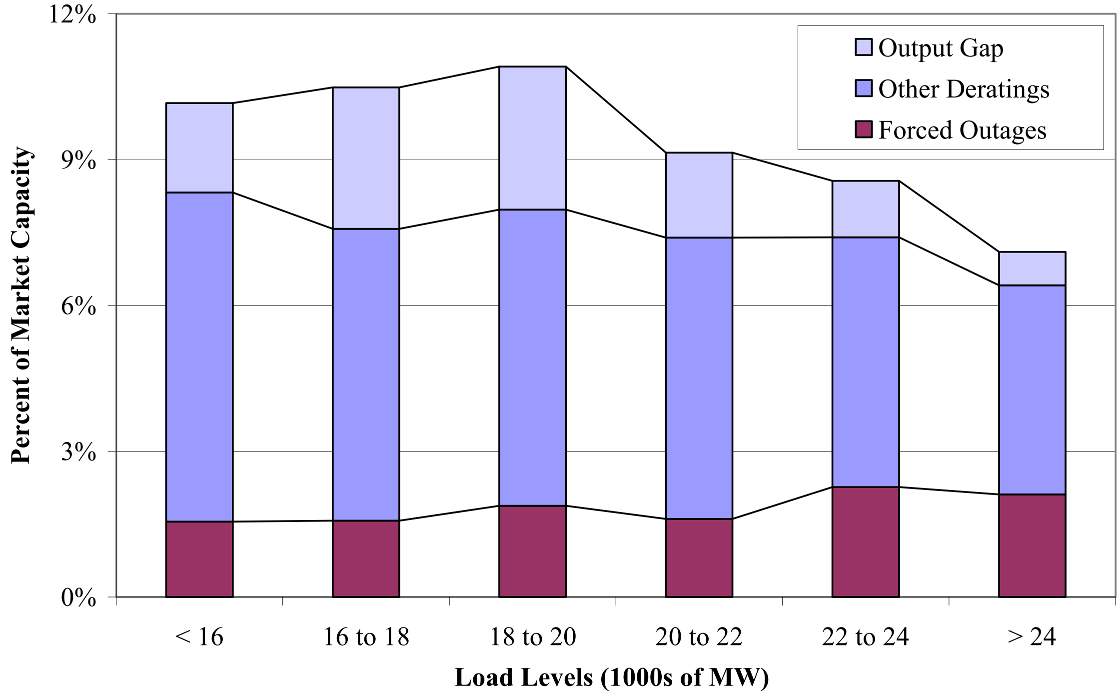
Economic withholding is measured with a statistic known as the "output gap", computed as the difference between a unit's capacity that is economic at the prevailing Energy Clearing Price ("ECP") and the capacity of the unit that is actually supplied into the energy market. To measure this, actual supply was compared to an estimate of economic capacity that is based on a proxy for each unit's marginal cost (i.e., a competitive benchmark for each unit's bids). The primary competitive benchmark used in this study is a "reference price" that is based on past accepted bids from each resource, although the output gap analysis is also conducted using variable production costs as the benchmark.

The analysis of physical withholding focuses on the total derating level for each unit, computed as the difference between a unit's maximum capability and the current rating (high operating limit) for the unit. The derating quantities analyzed in this report exclude planned outages and long-term forced outages because they are much less likely to constitute strategic physical withholding and including them could mask true physical withholding.

The empirical evaluation of the output gap and physical withholding provide strong evidence that the New England wholesale market is workably competitive. The primary empirical evidence supporting this conclusion is: a) the declining levels of the output gap and deratings that occur as load increases (i.e., in periods when the exercise of market power is most likely) and b) the lower levels of output gap and deratings for large participants. Measuring withholding as a percentage of market capacity, Figure E-2 shows the levels of both the output gap and total deratings at various load levels.

Total deratings is shown as the sum of its two component parts -- forced outages and other deratings. When a unit is on an outage, it is generally fully derated (i.e., its rating is zero for the hour so that its derating equals the unit's entire capability). The figure shows the marked decline in withholding as load reached the highest levels during 2001. These results and others in this report provide substantial evidence that the output gap and deratings did not include significant quantities of economic or physical withholding in periods when the slope of the supply curve was most steep – i.e., when the incentive to exercise market power was greatest.

Figure E-2
Total Potential Withholding by Load Level
January to December 2001



Source: ISO New England Operations and Market Settlements Databases. Potomac Economics analysis.

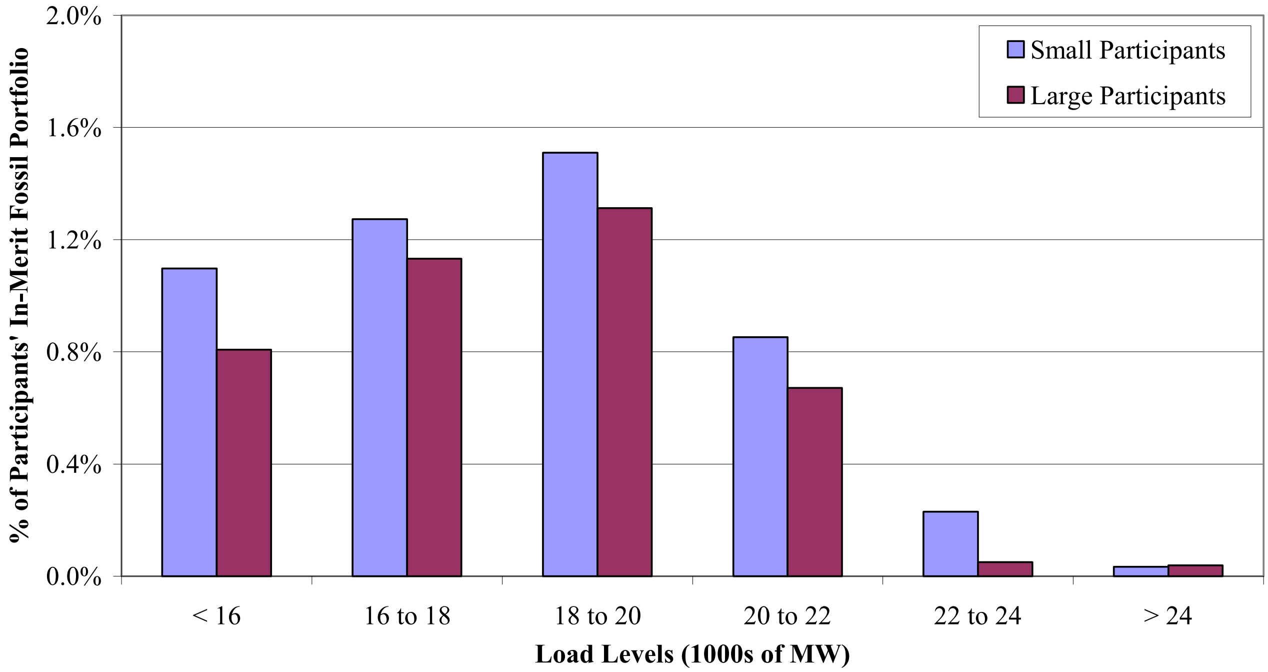
The empirical results also indicate that as a share of their portfolios large participants tend to exhibit smaller deratings and output gaps than smaller participants do. To illustrate this, the figures below compare results for large and small participants separately for the output gap (potential economic withholding) and total deratings (potential physical withholding).

Empirical Results of the Withholding Analysis

With regard to the output gap, a reliable comparative assessment must account for differences in the fuel mix of the suppliers. This can be accomplished by comparing the relative output gap of large participants (with available capacity exceeding 1200 MW) and small participants associated with their in-merit fossil-fired generation portfolio, shown in Figure E-3. This figure illustrates three important results of this analysis.

First, the output gaps for participants of all sizes are very small, averaging less than 1 percent of their portfolios. Second, the output gaps are smaller under the highest load conditions when the incentive to withhold is greatest. Third, the output gaps for large participants are less than those for small participants under all load conditions.

Figure E-3
Average Output Gap by Size of Participant During 2001
Fossil Units with Low Out-of-Merit Frequency



Source: ISO New England Operations and Market Settlements Databases. Potomac Economics analysis.

The analysis underlying Figure E-3 includes only fossil-fired generation, which constitutes roughly two-thirds of the capacity in New England. The output-gap comparison based on these resources only (excluding hydroelectric and nuclear capacity) is more reliable than those using broader measures. The output gaps computed for hydroelectric resources are less reliable because the opportunity costs facing these resources can vary substantially in a manner than is not accounted for in the reference price. Nuclear units generally produce at full output due to their low operating costs and the high costs of shutting down and starting back up, which results in an output gap of zero for these resources.

Units that are dispatched out-of-merit more than 20 percent of the hours that they run (roughly 25 percent of the total capacity) are excluded from the analysis behind the results in Figure E-3. This is justified because out-of-merit dispatch, which occurs in New England frequently as a result of NEPOOL's congestion management rules and procedures, changes the suppliers' bidding incentives. Units dispatched out-of-merit are paid their bid price (unless mitigated) rather than a market-clearing price. Under such "as-bid" compensation, suppliers lacking market power will raise their bid price to the expected market price in the constrained area.

Other findings of the output-gap analysis in this report relate to changes in the gap associated with out-of-merit dispatch and the implementation of three-part bidding. Comparing the average output gap associated with units that are generally in-merit when they are dispatched (out-of-merit dispatch less than 20% of total dispatch) with that of all units showed that the output gap for the in-merit units was roughly half as large as the average output gap for all units. This is consistent with the as-bid incentives facing units frequently dispatched out-of-merit.

However, suppliers that have locational market power in transmission-constrained areas have an incentive to raise their bid prices by more than the as-bid incentives would dictate. This form of market power is addressed by Market Rule 17 that employs bid screens to trigger mitigation of these out-of-merit bids, and it is not evaluated in this report.

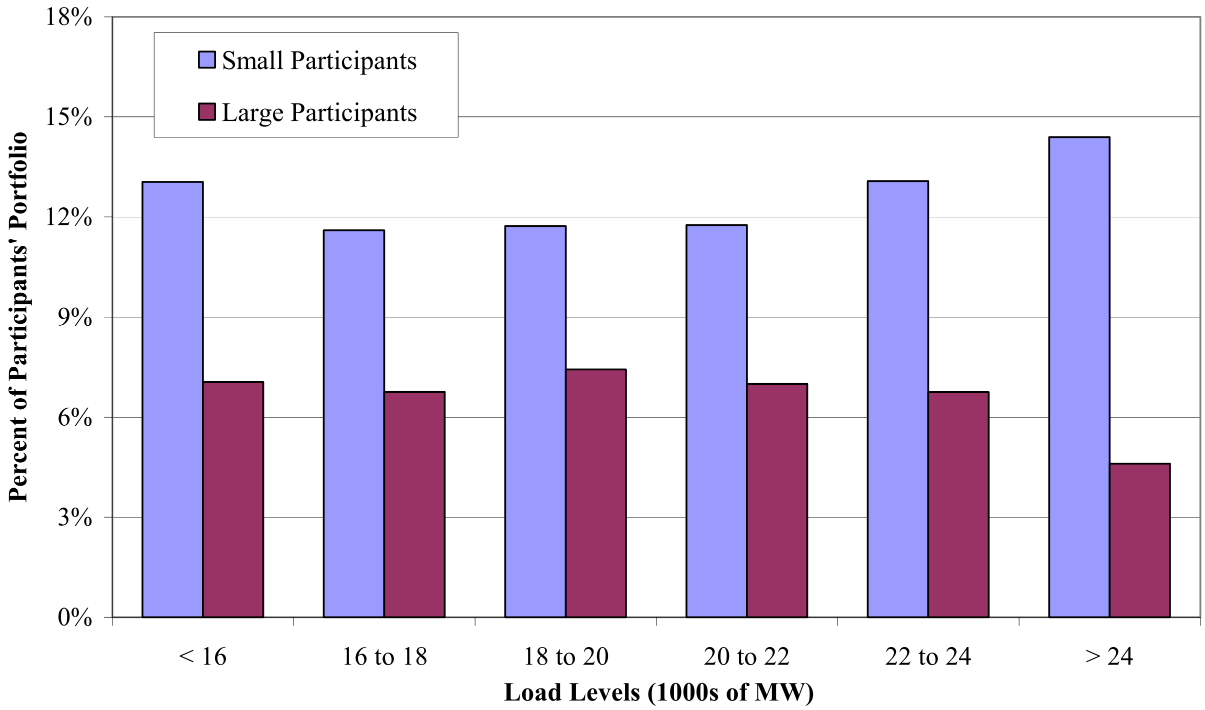
With regard to three-part bidding, the report evaluates the difference in bidding patterns before and after the implementation of the rule changes that allow suppliers to bid their start-up and no-load costs in addition to their energy bid. Three-part bidding was introduced on July 1, 2001. Because start-up and no-load costs are legitimate components of the marginal costs facing a generator, an owner who can submit only an energy bid may rationally increase its energy bid to prevent its unit from being committed and dispatched at a loss. This incentive can result in market inefficiencies and may explain a portion of the output gap that occurred before July 1.

Hence, the report compares the average output gaps for July to December 2000 versus July to December 2001. The average output gap in the 2001 timeframe was less than half of the comparable values in 2000. Although a number of other factors could have influenced these output gap differences, this result is consistent with the expectations described above for 3-part bidding.

Empirical Results of the Physical Withholding Analysis

The output-gap analysis assesses economic withholding while the following analysis of deratings is intended to reveal whether patterns exist that suggest strategic physical withholding has been a concern. Figure E-4 shows a comparison of total deratings for large and small participants at various load levels. Like the output gap results, the total deratings by large participants are consistently lower (as a share of portfolio capacity) than deratings by smaller participants.

Figure E-4
Total Deratings by Participant Size
January to December 2001



Source: ISO New England Operations and Market Settlements Databases. Potomac Economics Analysis.

The figure also shows that the percentage of large participants' portfolios that is derated decreases as demand grows to the super-peak levels when they would have the largest incentive to withhold. In contrast, the percentage of deratings by smaller participants is highest at the super-peak levels. These results again support the conclusion that the market in New England during 2001 was workably competitive.

While the analysis shows that the New England market has not been subject to systematic withholding, it cannot exclude the possibility that discrete instances of physical withholding occurred via specific outages or deratings. Therefore, ISO New England's physical-audit program designed to verify that outages and significant deratings are legitimate remains an important program to detect and deter this conduct.

Results of the Econometric Analysis

While the factors that may determine whether a supplier is economically or physically withholding resources can be individually analyzed as described above, econometric tools allow a fuller analysis of the relationship of each factor to the output gap or derating quantities while accounting for the other factors' effects. The factors encompassed in the econometric analysis include the strategic factors (participant size and peak demand conditions) as well as non-strategic factors that can help explain the output gap and deratings (average age of the portfolio, portfolio fuel-type shares, seasons, fuel prices, hour type, and others).

The results of this analysis show that neither the output gap nor the deratings are higher for a) large participants or b) under super-peak demand conditions (i.e., the highest 1 percent of hours). Further, the econometric analysis combines these two strategic variables to determine whether potential physical or economic withholding increases for *large participants under peak demand conditions*. Like the other results, the analysis does not indicate that the output gap or deratings increase in this case. These results are consistent with the descriptive analysis presented above and support the inference that the New England wholesale electricity market is workably competitive.

Analysis of High-Priced Hours

The report separately analyzes the highest-priced hours during the Summer 2001 to evaluate whether these prices efficiently reflected market conditions or were inflated by substantial withholding, flaws in the market rules or procedures, or unjustified actions of the ISO. Given the nature of supply and demand in the current wholesale spot electricity markets, super-peak conditions render prices far more sensitive to withholding or to market rules that do not facilitate full utilization of the system's resources. Since prices in these hours can be many times larger than the average price, the costs associated with unjustified price increases can be large even when such periods are relatively infrequent and short-lived.

I analyzed all hours when the ECP was greater than $200 per MWh, which included 18 hours during 2001. In 15 of these hours, the ECP was $1000 per MWh and was set by import transactions. The analysis seeks to determine whether these imports were economic and should have set the ECP in New England, and the extent to which the recently implemented pricing reforms will improve price determination under these conditions. The report also evaluates whether the New York ISO's transaction-scheduling rules and procedures contributed to inflated prices in New England in these hours. Finally, the report analyzes whether economic or physical withholding may have contributed to the high prices in these hours.

Table E-1 summarizes the analysis of the 18 high-priced hours. In the 15 hours that the ECP was $1000 per MWh, the price was set by an external transaction, and column (4) in the table shows the quantity of $1000 energy imports accepted by the ISO . These imports are accepted not because they are less expensive than all available internal energy resources, but rather to maintain the reserves in New England. Each MW of imported energy allows the ISO to create 1 MW of reserves internally to New England. If these reserves are worth the $1000 per MWh price tag, then the $1000 per MWh energy price is legitimate and efficient.

Table E-1
New England External Transactions in High-Priced Hours
Hours with ECP > $200 During Summer 2001

Date and Time (1)	Energy Clearing Price (2)	Reserve Shortfall (3)	Imports Accepted @ $1000 (4)	Excess Available Reserves (5)	New York Hour Ahead Price (6)	Forgone Economic Imports from NY (7)	Total Deratings (8)	Output Gap (9)
7/23/01 - 6 PM	$1,000	0	288	553	$94	0	14%	0.3%
7/23/01 - 7 PM	$1,000	0	288	1,217	$86	0	14%	0.4%
7/24/01 - 10 AM	$226	0	0	1,049	$212	0	11%	0.1%
7/24/01 - 1 PM	$1,000	-53	321	0	$1,000	150	11%	0.0%
7/24/01 - 2 PM	$1,000	-121	334	0	$1,000	350	11%	0.0%
7/24/01 - 3 PM	$1,000	-38	352	0	$999	400	11%	0.0%
7/25/01 - 12 AM	$1,000	0	352	199	$846	0	10%	0.0%
7/25/01 - 1 PM	$1,000	-12	352	0	$1,000	0	9%	0.0%
7/25/01 - 2 PM	$1,000	-301	352	0	$1,000	0	8%	0.0%
7/25/01 - 3 PM	$1,000	-391	352	75	$5,329	0	7%	0.0%
7/25/01 - 4 PM	$1,000	-115	352	15	$1,156	0	8%	0.0%
7/25/01 - 5 PM	$1,000	-289	352	299	$1,064	0	10%	0.0%
7/25/01 - 6 PM	$1,000	0	352	487	N/A	0	10%	0.3%
7/25/01 - 7 PM	$1,000	0	352	752	N/A	0	12%	0.9%
8/9/01 - 12 AM	$1,000	0	352	358	$999	0	6%	0.0%
8/9/01 - 1 PM	$1,000	-249	33	53	$1,000	0	6%	0.0%
8/9/01 - 3 PM	$243	-1,120	0	0	N/A	0	8%	0.1%
8/9/01 - 4 PM	$217	-847	0	76	N/A	0	8%	0.1%

Source: ISO-NE Transactions and Market Data, NYISO Transaction Bid Data, Potomac Economics analysis.

In most cases, imports at this price level are only economic if the ISO would be reserve-deficient without them. Therefore, column (3) shows the amount by which the ISO fell short of reserves in the hour. This shortfall in reserves is then compared to the excess available reserves (resources that are available to provide reserves that have not been designated for energy or reserves) shown in column (5). This comparison shows that in four hours in which the ECP was set at $1000 per MWh (July 23 at 6 pm and 7 pm and July 25 at 6 pm and 7 pm), there appeared to be sufficient reserves within New England to satisfy its requirements without the out-of-merit imports.

The report does not conclude from these data that the ISO erred in accepting the out-of-merit imports. First, the transactions are scheduled 30 minutes prior to the hour and must cover the anticipated needs over the entire hour – up to 90 minutes from the time the external transaction is selected. The pricing reforms recently implemented by the ISO address this uncertainty by changing the market rules to allow the out-of-merit imports to set prices only when they are truly needed to meet either the energy or reserve obligations of the ISO. Second, due to data limitations it is not possible to verify that all of the resources indicated as available to provide reserves actually could have been selected by the ISO. They may have been unavailable due to physical limitations (e.g., response-rate limitations, minimum down-time restrictions, or transmission constraints) or economic considerations (e.g., high energy or reserve bid prices, high opportunity costs, or long minimum run-times).

The table also shows how the New York ISO's scheduling process contributed to the high prices in New England. Exports from New York are scheduled only when the bid price for the export is greater than the hour-ahead price (transactions are scheduled hourly). The hour-ahead price is used for scheduling purposes only while settlements are based on the real-time prices produced each five minutes. Table E-1 shows that, with the exception of July 23, for each hour that the New England ECP was $1000, the New York hour-ahead price was close to or greater than $1000 per MWh (the N/A indicates that the New York market was in shortage and no clearing price was possible). When this occurs, no exports priced lower than $1000 per MWh will be made available to New England, even when prices in the real-time market are substantially lower.

Column (7) shows that there were three hours when exports that would have been scheduled from New York based on New York's real-time price failed to be scheduled due to the hour-ahead price. In two of the three hours, the quantities were large enough to eliminate the need for ISO New England to accept any $1000 per MWh transactions. Fortunately, the market-rule inconsistencies that cause the hour-ahead price sometimes substantially to exceed the real-time price are being addressed by the New York ISO prior to this summer.

Finally, Table E-1 shows the extent to which economic or physical withholding may have contributed to the high prices. Column (9) shows the output gap in each hour, and column (8) shows the quantity of deratings in each hour. The output-gap results shown in this table indicate that economic withholding did not play an important role in setting these prices. For purposes of this table, I computed the output gap assuming that the ECP was $950 to ensure that any internal units raising their energy bid close to $1000 per MWh and contributing to the ISO's decision to accept $1000 imports would be detected.

The table shows that in these high-priced hours the deratings ranged from 6 percent to 14 percent of the market capacity and the average was approximately 8 percent. The highest deratings occurred late in the day on July 23 and 25th. I examined the increase in deratings that occurred in those hours and found that the increases were primarily due to deratings on pump storage or other hydroelectric resources, which is consistent with the need to manage the limited available production capability of these units.

Although the quantities of deratings in most hours were not anomalously large, it is not possible to conclude that none of the deratings in these hours reflected strategic physical withholding. Therefore, the ISO's program of physical audits of the deratings of generating resources remains an important element of the monitoring program to detect and deter physical withholding.

This analysis supports a number of conclusions regarding these high-priced hours. First, in the majority of the high-priced hours, the ISO was sufficiently deficient of internal resources that acceptance of the $1000 per MWh imports was warranted. However, this deficiency may not have prevailed for the entire hour so that the $1000 per MWh price may have been justified for only a portion of some of these hours. The recently implemented pricing reforms will address this issue, as well as cases in which the ISO accepts an out-of-merit import uneconomically due to uncertainty regarding its need for the import at the time that it is accepted.

Further, if the reserves that the ISO was seeking to maintain had a value of less than $1000 per MW, the acceptance of the imports and associated energy price may not have been economic. Under the current market rules, the reserve requirement is absolute so that this value is presumed. Over the longer term, however, NEPOOL may consider implementing a demand curve for reserves, establishing an explicit value for reserves that would govern the ISO's actions to maintain the reserves and would determine the resulting energy prices. Such a demand curve would reflect the increasing marginal value of reserves as the quantity of reserves falls (and the deficiency grows).

Second, the New York ISO market rules related to scheduled exports from New York to New England did restrict New England's access to lower-cost imports in a few hours. If unaddressed, these conditions would likely reoccur under peak-demand conditions this summer. However, the New York ISO is implementing changes to its market models that should minimize this possibility in the future. This issue should continue to be monitored to ensure that these changes are effective.

Last, no clear evidence of economic or physical withholding during these high-priced hours emerged from this analysis. However, it is important to continue to monitor for such behavior, particularly in the peak-demand hours when the exercise of market power is most likely. This monitoring should include the types of screening and analysis of withholding presented in this report and, in the case of physical withholding, should be complemented by random physical audits to verify the technical justifications accompanying forced outages and significant deratings.

II. Introduction

The ISO New England's wholesale electricity markets that began operating in May of 1999 were designed to facilitate wholesale competition in the region, which promises substantial economic benefits. The regulatory reforms that ultimately created these markets, beginning with the Energy Policy Act of 1992 and continued by the Federal Energy Regulatory Commission ("FERC"), were intended to achieve economic benefits by allowing the market to guide short-term and long-term production and consumption decisions.

In the short-term, the market should minimize the production costs of meeting demand and set prices equal to the marginal value of electricity. In the long-run, the market should promote efficient investment, retirement, and demand-side decisions. Because these short and long-term benefits are contingent on workable competition, assessing the competitive performance of the market is a key component of the market monitoring function and fundamental to successful industry restructuring.

The competitive performance of the market depends both on the efficiency of the market rules and design, as well as the competitive structure of the market. These factors together create the incentives to which the market participants respond. The ISO New England Board of Directors requested two reports to assess these factors.

My first report assessing the New England market focused on the efficiency of the market rules and procedures by examining pricing during the peak periods in the Summer 2001 (hereinafter "Pricing Report").[1] That report found a number of flaws in the market rules that generally caused peak prices to be understated. In response to the conclusions and recommendations in that report, the ISO New England recently filed a number of market reforms with the Commission.[2]

This report will assess the competitive performance of the market by evaluating the degree to which the conduct of market participants is consistent with workable competition. The report draws on well-accepted economic principles related to imperfectly competitive markets to develop a theoretical model of electricity supply competition. This model is used to establish those market conditions and other factors that should be correlated with market power abuses (if significant market power exists), including market demand levels and participant size. Using the

New England bid data for 2001, empirical tests can be developed consistent with this model to determine whether market power has been a substantial issue in the New England markets.

The report seeks to identify attempts to exercise market power by withholding resources to increase energy market prices. The empirical tests described above are intended to differentiate anticompetitive withholding of resources from conduct that is competitively justified.

This task is complicated by two factors. First, imperfections in the market rules can create incentives for market participants to engage in conduct that may appear to be anticompetitive withholding. Second, the single-part bidding structure that prevailed prior to July 1, 2001 may have caused the bids for some resources to be justifiably adjusted to account for start-up and no-load costs to prevent units from being committed uneconomically. The results presented in this report seek to account for both factors.

The analysis in this competitive assessment complements the analysis in the report recently produced by Bushnell and Saravia.[3] The Bushnell and Saravia study utilized a simulation model to estimate a competitive benchmark for prices in New England assuming the marginal generating costs for fossil units equal their estimated variable production costs. This competitive price benchmark is then compared to actual prices in New England to assess the competitiveness of the New England markets. Alternatively, the analysis in this report seeks to directly assess the extent to which suppliers have engaged in economic or physical withholding in an attempt to raise market-clearing prices.

In addition to this analysis of withholding, the report analyzes in detail the highest-priced hours during the summer of 2001. This analysis seeks to determine the extent to which withholding by participants, inefficient market rules and procedures, or unjustified actions by the ISO may have contributed to inflated price levels in these hours. Like the Pricing Report, this section of the report evaluates whether the market provided efficient price signals in these hours.

A. Wholesale Electricity Markets

Restructured wholesale electricity markets generally use a clearing-price auction to efficiently dispatch the generation to meet the energy and ancillary services demand in real time, generally referred to as "spot" markets. Some of these auctions recognize the constraints imposed on the delivery of electricity by the transmission network by setting clearing prices that vary by location (generally referred to as locational marginal prices or LMP). New England's current market employs a uniform clearing-price auction, setting a single clearing price and dispatching generation out-of-merit order when necessary to resolve transmission constraints on the network. The supply in New England is dispatched every five minutes to meet the real-time load and to determine the 5-minute Real-Time Marginal Price ("RTMP"). The time-weighted average[4] of these 5-minute prices over each hour, called the Energy Clearing Price ("ECP"), is used to settle the energy market for that hour.[5]

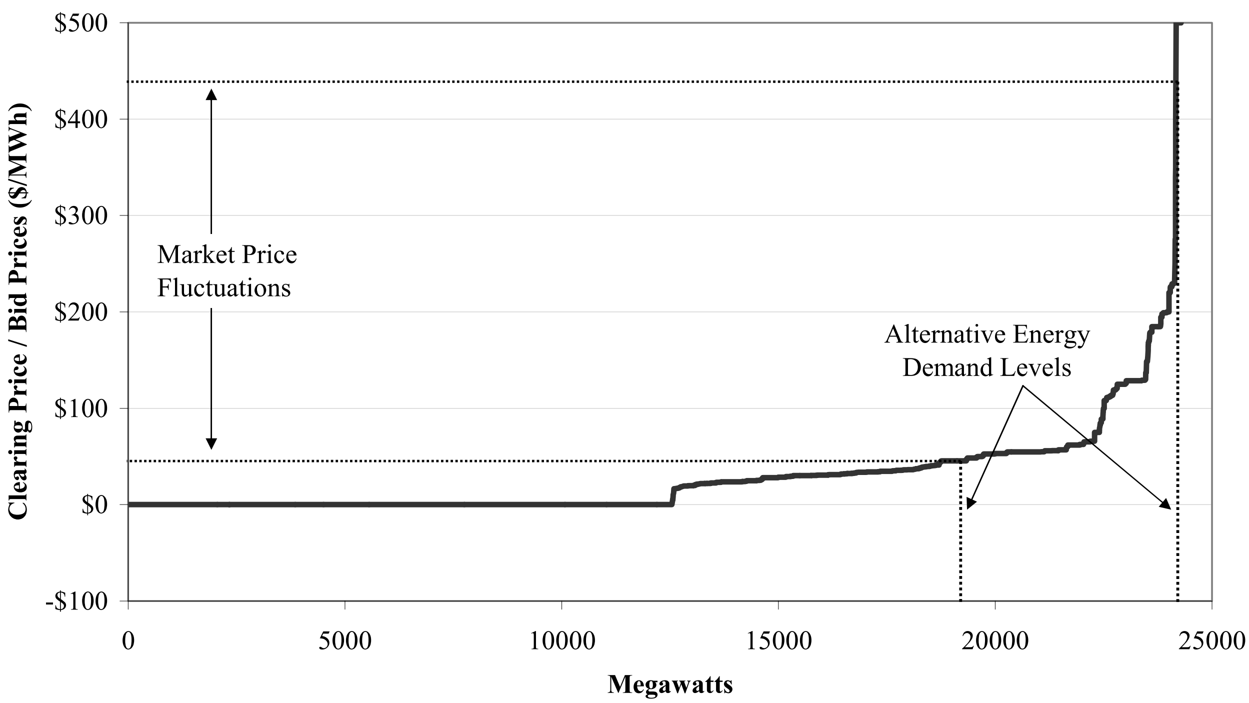
Although a small fraction of the power consumed in New England settles through the spot energy market, the spot market should guide the dispatch of all generation within the region. A substantial portion of the generation in the region is sold through bilateral energy contracts. Nonetheless, the buyers and sellers under bilateral contracts should still respond to spot market prices since suppliers can meet their bilateral obligation with spot market purchases when the spot energy price is less than the marginal cost of producing from their own generation. For example, assume a generator with a $40/MWh marginal that has signed a bilateral energy contract at $50/MWh. In an hour where the spot price is $20/MWh, the generator would make a $10/MWh profit supplying the bilateral obligation from its own generation ($50-$40) versus a $30/MWh profit turning its unit off and supplying the obligation with energy purchased in the spot market ($50-$20). Therefore, the spot market should guide the dispatch of all generation in the region.

Under New England's proposed market reforms in moving to its "Standard Market Design" ("SMD"), it will establish day-ahead and real-time energy markets that will set location-specific clearing prices to recognize transmission constraints in the region. This approach is consistent with FERC's proposed "Standardized Transmission Service and Wholesale Electric Market Design", as well as the current energy markets in New York and the Mid-Atlantic.[6]

Since there is generally no economically viable method for storing large quantities of electricity, generation must satisfy demand in real time. Consequently, market prices will fluctuate to clear the market in real-time as demand and supply conditions change.[7] When demand is high or there are some large generating units unavailable, higher-cost generating units must be run and spot market prices will rise. Alternatively, spot market prices fall when demand is lower and can be met by lower-cost generating units. Just as the market prices fluctuate with supply and demand conditions, the competitiveness of the market can vary as supply and demand conditions vary.

Figure 1 shows the price dynamic in a competitive spot electricity market. This figure shows the New England supply curve corresponding to the 6:00 pm hour on August 9, 2001. This supply curve was chosen for illustrative purposes to show how prices fluctuate as demand rises and falls. Under peak demand conditions, prices can rise in percentage terms by amounts seldom seen in other product markets (e.g. prices frequently rise from an average of less than $40 per MWh to more than $1000 per MWh).

Figure 1
Price Fluctuations in Restructured Electricity Markets
Illustrative Supply Curve - August 9, 2001



It is important to understand this dynamic for two reasons. First, considerable price fluctuations can occur in competitive electricity markets that do not have substantial excess supply. Thus, one cannot assume that such fluctuations are the result of market power. Further, policies that would restrict competitively justified fluctuations in the name of market power mitigation will undermine the efficiency of the competitive market. Therefore, it is essential to accurately differentiate between competitive market outcomes and price increases due to market power.

Second, this dynamic is related to the existence of market power. As described in more detail below, the incentive to exercise market power is dependent on the sensitivity of the market price to changes in supply. The supply curve figure shows that prices become much more sensitive as demand approaches the relatively steeply-sloped portion of the supply curve (toward the right of the figure).

Given a market where all suppliers are paid the clearing price, suppliers lacking market power will have the economic incentive to produce when the cost of producing additional output is less than the clearing price. Understanding the incentives of suppliers in restructured electricity markets is essential for assessing whether suppliers are attempting to exercise market power and, more broadly, for evaluating the overall competitiveness of the market.

B. Market Power in Electricity Markets

Market power is defined by most economists as the ability of a firm to profitably raise market prices substantially above competitive levels. It must be understood that market power exists in virtually every product market – only perfectly competitive markets lack market power entirely. It is generally impossible or very costly to eliminate all market power, which explains why economists generally employ a standard of "workable" competition rather than perfect competition in assessing the performance of a market. Workable competition allows for the fact that market power may exist in a market at a level below levels that would raise public policy concerns and would be uneconomic to attempt to eliminate.

Market power is generally exercised in electricity markets by withholding supplies from the market in an attempt to raise the market-clearing price. Withholding of supplies may take two forms. First, a supplier can physically withhold capacity from the market by claiming a generation outage that is not technically justified or by simply not offering the resource into the market when it would be economic for the supplier to do so. Any resource that is running, whether bid flexibly into the spot market or self-scheduled, is not physically withheld.

Second, a supplier can economically withhold the unit by bidding it at a price that is higher than the unit's marginal cost in order to reduce the resource's output and raise the market price. A generator's marginal cost is the incremental cost of producing additional output, including opportunity costs and incremental risks associated with unit outages. For a large share of the fossil resources in most electricity markets, the units' marginal cost equals their variable production costs. However, units with energy limitations such as hydroelectric units that must forego revenue in a future period when they produce in the current period incur an opportunity cost associated with producing that can cause their marginal costs to be much larger than their variable production costs.

These two types of withholding are consistent with the notion that both capacity and bid price are strategic variables in electricity markets. Physical withholding corresponds to the use of capacity as a strategic variable, while economic withholding corresponds to the use of price as a strategic variable.[8] As explained below, a good way to model supplier behavior in electricity markets is to use techniques which recognize both price and capacity are strategic variables within the control of the generator.

1. Previous Market Power Studies

Measuring market power in electricity markets has received increasing attention recently. This is in part because experience with restructured markets has developed to the extent that sufficient data is available to make empirical inferences. Market power studies have also proliferated as a result of the high-profile events associated with California's restructuring experience. Indeed, some of the most detailed recent studies have focused on California's markets.[9]

Market power studies in restructured electricity markets are of two general categories. The first category includes simulation analyses that estimate a competitive equilibrium level of prices that would prevail if all suppliers acted in a competitive manner.[10] The Bushnell and Saravia (2002) study of the New England market is an example of this type of study. These analyses generally use estimates of each generator's variable production cost as proxies for the generator's marginal costs. Having estimated a competitive equilibrium level of prices, these studies then generally compute a "mark-up" by computing the average difference between the actual clearing prices that prevailed in the market and this estimated level.

This mark-up statistic can be a useful diagnostic index for the market in two respects. First, changes in this index over time can provide useful information for assessing the competitive performance of the market. Second, the relative differences in the index between markets with similar attributes can also provide useful insight.

However, it is not appropriate to interpret this mark-up index as a measure of market power that has been exercised in the market for a number of reasons. First, the estimated mark-up can be due to market rules that cause the incentives facing generators or the determination of market prices to differ from the assumption made in the simulation analysis. Second, since the true marginal costs of some generators can substantially exceed their variable production costs, the estimated competitive equilibrium level may be understated causing the mark-up to be overstated (typically under peak demand conditions).

Lastly, the simulation analyses do not usually account for the unit commitment process that determines which units to turn on for the following day. By assuming all available units (i.e., that are not on outage) are on-line in the simulation, the competitive equilibrium price will again tend to be understated and the mark-up overstated (typically under off-peak demand conditions). However, Bushnell and Saravia account for this in the New England study by computing the mark-up in one scenario by comparing prices estimated with the actual bids versus prices estimated with variable production costs. Both estimates assume that all units have been committed so the estimated mark-up would not be systematically overstated due to the difference in the assumed commitment of generation.

These factors do not diminish the usefulness of the mark-up index as a diagnostic statistic that can be used to evaluate the performance of the market. They do, however, preclude the mark-up from being interpreted as a reliable measure of the market power that exists in the market.

Withholding analyses make up the second category of market power studies. These studies focus more directly on the behavior of market participants by seeking to determine whether capacity has been strategically withheld from the market.[11] This report falls into this second category of market power studies. The key analytic task in these studies is to differentiate strategic withholding from competitive conduct that could appear to be withholding. For example, a forced outage of a generating unit may either be legitimate and expected in competitive electric markets, or a strategic and deliberate attempt to raise prices by physically withholding the unit.

To differentiate between these two alternatives, the empirical analysis in this report is guided by economic theory that determines the market conditions and participant characteristics that would tend to create the ability and incentive to exercise market power. This economic theory and the associated market power hypotheses to be analyzed in this report are described in the next section.

2. Incentive to Withhold Resources

A large body of economic theory has been developed that indicates how firms behave in markets that are not perfectly competitive. This theory is premised on the assumption that firms seek to maximize their profits. In general, firms that lack market power will maximize their profits by offering all of their economic resources in the market (i.e., not withhold resources). Alternatively, firms that can profitably influence prices will withhold resources when they have the incentive to do so.

In New England, firms offer schedules of quantities and prices indicating how much energy they are willing to generate at any given price level – sometimes referred to as a supply function. Given that the bid prices and quantities are both strategic variables that firms may use to maximize their profits, the concept of supply function equilibrium developed by Klemperer and Meyer can be applied.[12]

Under this framework, each firm offers a supply function to maximize its profits, given that other firms also are offering supply functions to maximize their profits.[13] Each firm also recognizes that the supply function offered by its rivals will have an impact on its own profits because of the other firms' impact on the market clearing price. The equilibrium set of supply functions are those where all firms are maximizing their individual profits, yet no firm can do better by deviating from its individual supply function.

Klemperer and Meyer derived the general characteristics of the supply function equilibrium and others have also used the supply function equilibrium framework to evaluate the competitive incentives of firms in restructured electric markets.[14]

The basic result derived using this framework is the amount of output that will be produced by the supplier seeking to maximize its profit as a function of the key market conditions. This level of output, q^*, is given by:

(1) $$q^* = (p - MC(q^*)) / (dp/dS).$$

This represents the first order condition for each supplier and the derivation of this result is provided in Appendix A. In (1), p is the market clearing price and MC(q) is the marginal cost of the most expensive block of power offered by the supplier.[15] dp/dS represents the change in price ("dp") corresponding to a change in supply ("dS") – or the sensitivity of prices to changes in supply. It is easy to see that at some point under high demand that this value becomes very large.[16]

If the amount supplied given by (1) is compared to what would be supplied under competitive conditions, then the profit-maximizing level of withholding can be determined. Recall the competitive supply function is given by MC(q) = p (i.e., each firm increases its production until the marginal cost of its most expensive resource equals the market price). We will denote the quantity that solves the competitive supply function as q^C. The equilibrium amount withheld by a suppler is $w^* = q^C - q^*$. A firm would never offer to supply more than q^C because operating losses would ensue.[17] Hence w is constrained to be non-negative (i.e., negative values from the following formula mean that the optimal w is zero). Using (1), w^* becomes:

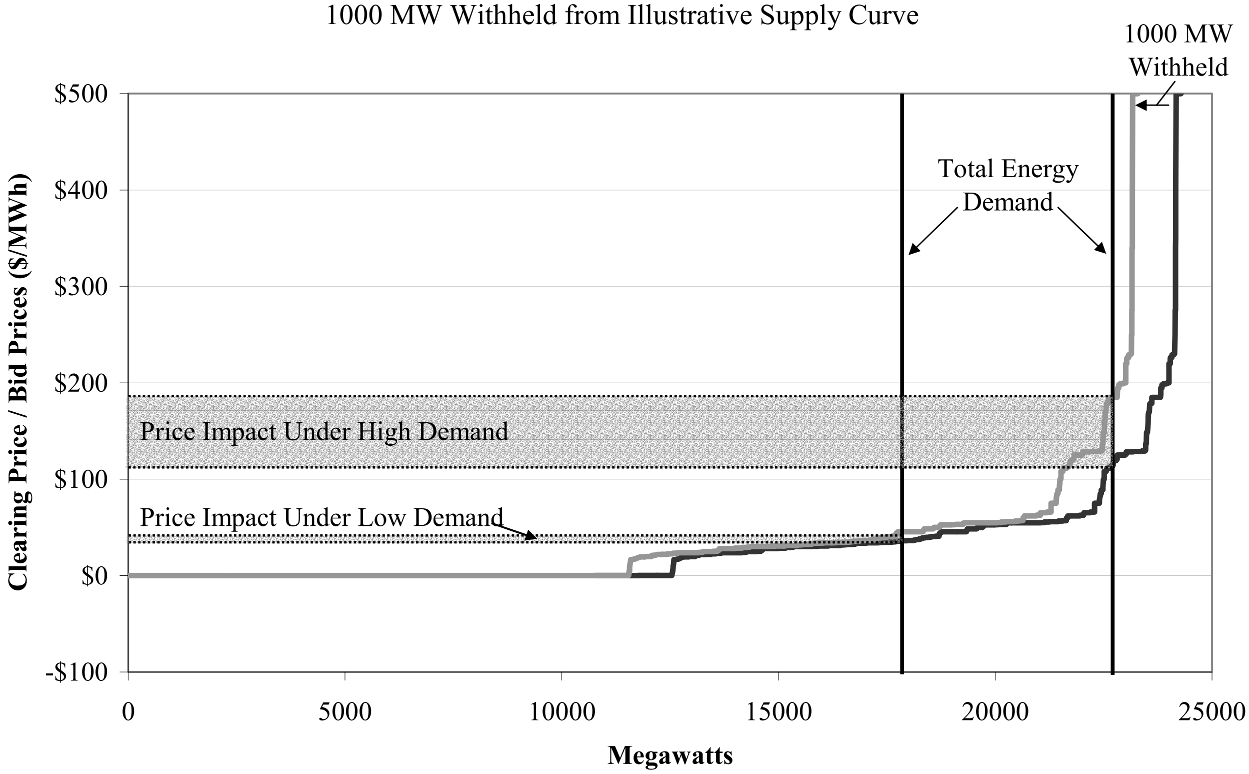
(2) $w^* = q^C - [p - MC(q^*)] / (dp/dS)$.

This equation shows that one can define conditions under which withholding is rational versus those where withholding is irrational (i.e., $w^* = 0$). Withholding will depend primarily on two factors. First, withholding will depend on the level of q^c. Assuming the portfolio of generating units are similar from firm to firm, q^c should be directly correlated with the size of the firm.

Second, withholding will depend on the sensitivity of prices to supply shifts (dp/dS), which defines the slope of the supply curve. Using the illustrative supply curve first introduced above in Figure 1, one can see when the slope of the supply curve is so important. Figure 2 shows how prices respond to 1000 MW of withholding when demand is at moderate levels versus at peak levels.

Figure 2 shows the shift in the supply function to the left that would occur if 1000 MW of low-cost resources were withheld from the market. Under peak conditions when the market is clearing on the relatively steeply sloped portion of the supply curve, the resulting impact on prices from the withholding is much larger than the price effects under moderate load conditions. This difference in price effects plays a key role in determining the profitability and, thus, the incentive to withhold. Fortunately, the vast majority of hours exhibit the load levels that correspond to the relatively flat portion of the supply curve where there is very little incentive to withhold.

Figure 2
Impacts of Withholding Under High and Low Demand Conditions
1000 MW Withheld from Illustrative Supply Curve



The intuition regarding the importance of firm size and the price elasticity is straightforward. Withholding is most likely to be profitable when lost profits from withholding are more than offset by the additional profits earned from the remaining sales made at the inflated price level. The amount of the remaining sales is determined by the firm's size and the increased profit on those sales, which depends on the slope of the supply curve.

To show how these two factors affect suppliers' incentives to withhold, I have constructed an example shown in Table 1, which is based on equation 2 above. This table shows how suppliers' optimal withholding amounts change as load increases and prices become more sensitive to changes in supply. This relationship is shown for two cases that vary by the size of the supplier. The smaller supplier case is assumed to have 500 MW of economic capacity versus 1000 MW of capacity assumed for the larger supplier.

It is important to remember some of the key assumptions made in deriving equation 2. First, demand is assumed to be unresponsive to changes in price. Price-responsive demand would reduce the incentive to withhold by making prices less sensitive to withholding. Second,

Table 1
Effects of Supplier Size and Price Sensitivity on Suppliers' Optimal Withholding Amount

	Load Level (MW)	Optimal Withholding (MW)	Withholding Percent (%)	Assumed Margin (MW)	dp/dS* ($/MW)
Smaller Supplier Case - Economic Quantity = 500 MW	16000	0**	0%	2	0.0017
	18000	0**	0%	5	0.0055
	20000	0**	0%	6	0.0062
	22000	0**	0%	30	0.0433
	24000	132	26%	500	1.3587
Larger Supplier Case - Economic Quantity = 1000 MW	16000	0**	0%	2	0.0017
	18000	0**	0%	5	0.0055
	20000	35	4%	6	0.0062
	22000	307	31%	30	0.0433
	24000	632	63%	500	1.3587

* Estimated based on the New England supply curve for August 9, 2001 shown in Figure 1.

** Withholding can only be positive -- the results of the formula in these cells produce values ranging from -676 to -91 MW in these six cases.

all suppliers are assumed to be the same size and react to one another. Therefore, the change from 500 MW to 1000 MW for all suppliers in the market will produce a bigger change in a supplier's optimal withholding than will the increase in size of a single supplier. Nonetheless, the example is useful in illustrating the importance of these two factors.

This example is based on equation 2. One of the key components of this equation is the supplier's margin – the difference between the market price and the supplier's marginal cost ($p - MC(q^*)$). The example shown in Table 1 assumes margins that increase as prices become more sensitive to withholding. This is logical because this margin should increase as the total impact of the joint withholding by all suppliers rises under peak conditions.

Given these assumptions, both cases show that the suppliers have no incentive to withhold output at moderate load levels when the supply curve is relatively flat. In the smaller supplier case, suppliers' optimal withholding is zero until super-peak load levels over 24,000 where suppliers should withhold 24 percent of their economic output. On the other hand, suppliers in the large supplier case would begin withholding when demand exceeds 20,000 MW, rising from 4 percent of the economic capacity withheld at 20,000 MW to 63 percent withheld at 24,000 MW.

This example illustrates the theoretical implications of the supply function equilibrium results shown in equation 2 which is the basis for the two primary empirical hypotheses that guide our empirical analysis:

Empirical Hypothesis 1: the incentive to withhold increases during periods of high demand when prices are relatively sensitive to changes in supply -- *ceteris paribus*, withholding increases under high-demand conditions.

Empirical Hypothesis 2: the incentive to withhold will be greater in a firm with a larger supply portfolio -- *ceteris paribus,* withholding will be greater in larger firms.

These empirical hypotheses will be investigated and tested in the following sections that examine the two primary forms of withholding. Section III provides the analysis of economic withholding while Section IV provides an analysis of physical withholding.

III. Economic Withholding

The first class of conduct analyzed in this report is economic withholding. Economic withholding can be defined as withholding a resource by raising its bid to raise the market price above competitive levels. As this definition implies, any analysis of economic withholding must determine when the bid for a resource has been raised above the level that would be bid if it faced workable competition. Stated another way, the analysis must determine how bidding and resulting output of a resource would differ from a supplier behaving as a price-taker (i.e., a supplier lacking market power).

In a competitive clearing-price market where no supplier can influence the price, suppliers maximize their profits by accepting the clearing price when that price is higher than their marginal costs of producing and not producing when it is lower (i.e., behaving as a price taker). The competitive conduct implied by this assertion is that each supplier should be bidding its marginal costs into the auction market as described in the prior section. A generator's marginal cost is the incremental cost of producing additional output, which is usually composed primarily of the variable production costs of the unit. However, marginal costs also include opportunity costs and other types of incremental costs. Opportunity costs are incurred when a unit with limited production capability over some period of time produces in the current period and, in doing so, foregoes revenue in future periods. An example of an incremental cost that is not generally included in variable production cost is the incremental risk associated with unit outages. When a steam unit operates in a manner to produce at its maximum technical capability, increased pressures and other factors may increase the probability the unit will suffer a forced outage. This incremental risk, if associated with the highest segment of output on the unit, is a component of the marginal costs for this segment of output.

Establishing a proxy for the units' marginal costs as a competitive benchmark is a key component of the analysis that is described below. This is necessary to determine the quantity of output that is potentially economically withheld. The following subsection describes how potential economic withholding is estimated given a competitive benchmark for each generating resource. The second subsection describes how the competitive benchmarks for each resource are estimated. The balance of this section presents the results of the economic withholding analysis relative to a number of factors reflecting market conditions and rules, as well as participant characteristics.

A. Measuring Economic Withholding

Economic withholding is measured in this report by the estimation of an "output gap", which is defined as the difference between the unit's capacity that is economic at the prevailing ECP and amount that is actually produced by the unit. This measure was introduced by Joskow and Kahn in a recent analysis of market power in the California electricity market.[18]

In essence, the output gap shows the quantity of generation that is withheld from the market as a result of having submitted bids above competitive levels. Therefore, the output gap for any unit would generally equal:

$Q_i^{econ} - Q_i^{prod}$ when greater than zero, where:

Q_i^{econ} = Economic level of output for unit i given the current ECP and competitive bid for the unit.

Q_i^{prod} = Actual production of unit i.

Only positive values from this formula are included in output gap. In addition, some are dispatched at lower levels than their current bid would suggest due to transmission constraints or reserve considerations. Therefore, we adjust for output that is bid at prices below the ECP but which is not produced. Hence the output gap formula used for this report is the following:

$Q_i^{econ} - \max(Q_i^{prod}, Q_i^{bid})$ when greater than zero, where:

Q_i^{bid} = bid output level at the ECP.

By using the greater of the actual production level or the bid output level at the ECP, units that are dispatched down due to transmission constraints or other factors are excluded from the output gap. Two other adjustments are made to the output gap values to allow them to better reflect potential economic withholding.

First, resources that are selected to provide reserves are not included in the output gap. This adjustment is justified because price increases in the energy market are generally accompanied by shortages of operating reserves. Therefore, relatively high-cost resources that provide reserves are generally as valuable to the market supplying reserves as they would be supplying energy.

Second, resources that are not economic to commit given their start-up costs are excluded from the output gap. Prior to the implementation of three-part bids to allow for start-up cost and no-load bids, conduct that appears to be economic withholding could occur to prevent units that are not economic from being committed. Therefore, we used the competitive bid proxy for each unit to determine whether the unit would make enough profit at the actual ECP to recover its start-up costs.

The start-up cost data were obtained from the ISO New England's NX-12 database consisting of generating unit data gathered by the New England Power Pool prior to ISO operation. Prior to July 1, 2001, any resources that fail this commitment test are excluded from the output gap values for that day. This exclusion has a very small effect on the output gap results, reducing the output gap quantities for the period prior to July 1 by an average of 2.5 percent.

B. Competitive Benchmark

As described above, the competitive benchmark assumed for the analysis is a critical element of estimating the output gap because it determines the economic level of output for a unit at a given ECP. This section describes how this benchmark is developed for the analysis of economic withholding.

As described above, generators lacking market power in a competitive clearing-price market should bid their marginal cost of production, including opportunity costs and costs related to incremental risks associated with unit outages. Therefore, the competitive benchmark used to compute the output gap should approximate the marginal costs of the supplier.

The primary competitive benchmark used in this report is a reference price that is generally calculated based on the historic accepted bids for each unit. If withholding is not rational under most market conditions as described in the introduction, then using the historic in-merit accepted bids of the supplier should provide a reliable indicator of the competitive bid level for the unit. The study confirms the primary output gap results based on reference values by also estimating the output gap based on estimates of variable production costs.

Reference values are calculated for the entire output range of the unit (in 10 MW segments for most units) as the lower of the mean or median of the accepted bids for each respective output segment over the current season. When there is an insufficient number of accepted bids for a segment because the segment was generally self-scheduled, the reference price is estimated from the lowest quartile of ECP's in hours when the unit was self-scheduled in comparable periods (peak vs. off-peak). This serves as a reasonable proxy for the generator's marginal cost by identifying the lowest spot prices under which the self-scheduled resource is willing to run.

For resources that are rarely scheduled, the reference price is calculated based on the lowest quartile of all bids submitted for the output segment. This is a reasonable approach because a generator would have to economically withhold its resource in nearly all hours for this reference price to overestimate the resource's marginal cost. This methodology is described in more detail in Appendix B.

Some economists have used variable production costs as a proxy for marginal costs. This is justifiable for a large share of the generating resources whose incremental cost of additional output is dominated by the variable production costs (i.e., fuel, emissions, and variable operating and maintenance costs). However, some resources have marginal costs that substantially exceed variable production costs.

Units with output restrictions, such as hydroelectric and some fossil resources, may experience opportunity costs that are much higher than their variable production costs, which are very close to zero for many of these units. These opportunity costs relate to profits foregone in future periods from producing in the current period. Other resource blocks, such as the highest output levels on some steam units, may experience a higher probability of forced outages when operating in that range.[19] The expected losses of an outage times the incremental increase in the outage probability associated with producing in this range can generate very high marginal costs for these ranges.

Establishing competitive benchmarks that account for these factors is particularly important because many of these resources represent the highest-cost supplies that are dispatched to satisfy energy or reserve requirements in peak hours. Therefore, using variable production costs as the competitive benchmark will generally show larger quantities of potential economic withholding than may actually be occurring. Nevertheless, output gap results based on variable production costs are presented later in this section to confirm the results obtained with the reference prices.

C. Descriptive Analysis of the Output Gap

The analysis presented in this section evaluates the output gap results relative to various market conditions and participant characteristics. The objective of this analysis is to determine whether the output gap quantities increase when those factors prevail that can create the ability and incentive for one or more suppliers to exercise market power. Therefore, we are not attempting to explain the output gap, but instead are testing whether it varies in a manner consistent with an attempt to exercise market power.

1. Demand Level

The first analysis of the output gap relates to its correlation with the level of demand in the market. This is important because the incentive to withhold resources and raise prices, to the extent that it exists, should be greatest in the peak demand hours when prices are most responsive to changes in supply. As previously described, the incentive to withhold does not rise gradually as the demand rises. Rather, it is associated with the slope of the supply curve, which only rises substantially at the highest levels. A more steeply sloped supply curve results in higher sensitivity of prices to withholding assuming relatively inelastic demand.

Therefore, the load levels analyzed focus most heavily on the peak demand periods. For example, the three highest load levels shown in the following figures correspond to 1.7 percent, 0.7 percent, and 0.3 percent of the total hours during 2001. This distribution of hours to each of the load levels and the relationship of the load levels to the slope of the supply curve are shown in Table 2.

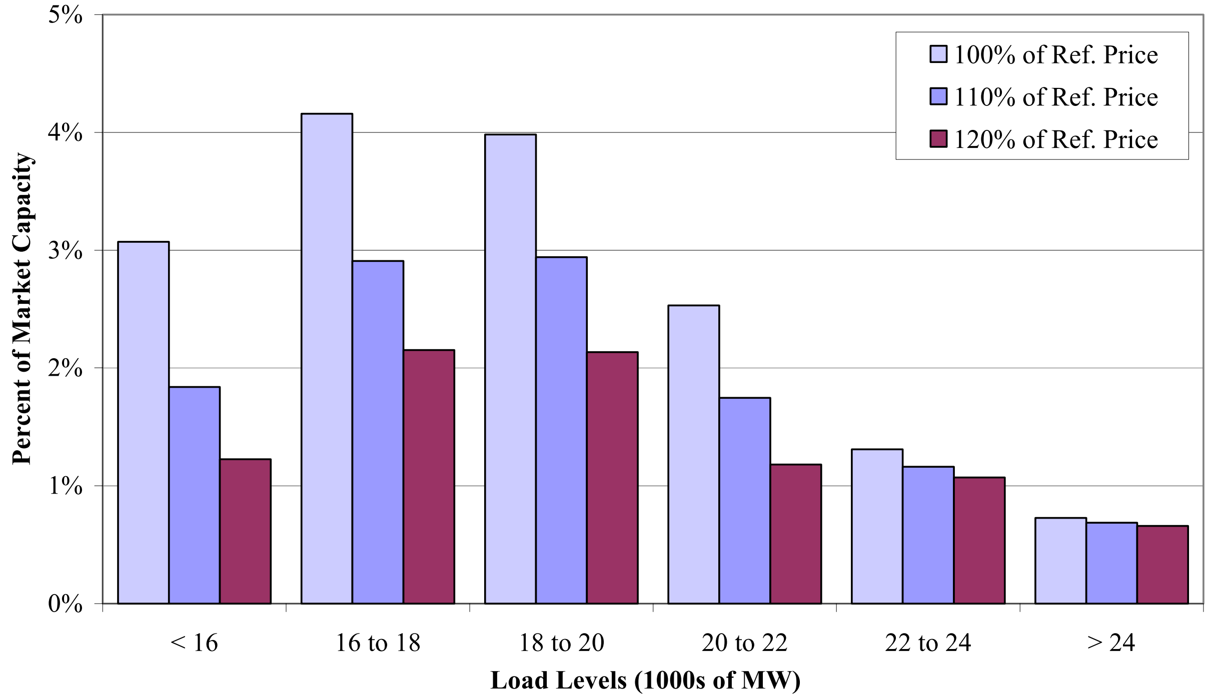
Table 2
Relationship of Load Levels to the Slope of the Supply Curve

Load Level	Percent of Hours	dp/dS*
> 16,000	69.2%	0.0017
16000 - 18,000	21.7%	0.0055
18000 - 20,000	6.3%	0.0062
20000 - 22,000	1.7%	0.0433
22000 - 24,000	0.7%	1.3587
< 24000	0.3%	> 3.0

* Estimated based on the New England supply curve for August 9, 2001 shown in Figure 1.

Figure 3 below shows how the output gap measured as a percentage of the total market capacity varies as the level of demand increases. An alternative approach would be to show the output gap as a percentage of the total economic capacity (capacity with marginal costs less than the ECP). Although this is a valid alternative method to measure the relative size of the output gap, it would tend to bias the results toward showing lower output-gap ratios in the highest load periods since these periods exhibit the highest ECPs and thus have the largest quantities of economic capacity. To avoid this potential bias in the results, therefore, each of the analyses presented in this report computes the output gap as a percentage of total capacity rather than economic capacity. This is true of both the market-level results as well as the participant-level results where the output gap is shown as a percentage of each participant's total capacity.

Figure 3
Average Output Gap by Load Level for All Units
Alternative Reference Price Cases: January to December 2001



Source: ISO New England Operations and Market Settlements Databases. Potomac Economics analysis.

The results shown in Figure 3 indicate a clear pattern of lower quantities of potential economic withholding at the highest demand levels. In addition, the estimated quantities at the highest demand levels are a very small portion of the total market capacity. This pattern is consistent with the hypothesis of workable competition.

The other noticeable change in output gap quantities is the increase that occurs between the lowest demand periods (< 16,000 MW) and the moderate demand periods (16,000 to 20,000 MW). This increase is likely the result of the fact that the lowest ECP's occur in lowest demand periods. When the ECP is relatively low, the corresponding quantity of resources that appear to be economic will also be very low. Since the figure shows the output gap as a percentage of the total market capacity (rather than the economic capacity), the output gap percentage will generally be lower when a relatively small share of the market's capacity is economic. Further, most of the resources that are economic when the ECP is very low are large baseload resources that generally bid at relatively low levels to ensure that they continue to operate at a constant level of output and avoid the start-up costs that would be incurred by shutting the unit down overnight.

In addition to showing the various demand levels, the figure shows three alternative reference price cases ranging from 100 percent to 120 percent of the reference prices. These cases vary in the level of the competitive benchmark used to determine the quantity that would be economic at the current ECP. For example, the 120 percent case would identify output from a unit as economic when the ECP is 20 percent higher than the unit's reference price.

The substantial reduction in the output gap that occurs between the 100 percent case and the other cases indicates that a large share of the output gap in the 100 percent case is associated with resources that have been bid at prices slightly above their reference price. This additional output gap quantity in the 100 percent case is not apparently consistent with a deliberate strategy to withhold a significant quantity of resources from the market to influence the market price. In contrast, it is likely indicative of the fact that any proxy for marginal cost will not account for short-term fluctuations in marginal costs that would result in concomitant fluctuations in bid prices. The fact that the larger differences occur at moderate demand levels when market power is much less likely to be an issue is consistent with this conclusion.

Therefore, the balance of the analysis presented in this report utilizes the 110 percent output gap case in an attempt to focus the analysis on patterns of potential withholding that may reflect an attempt to exercise market power, rather than a reflection of normal fluctuations in marginal costs.

The last observation that can be made from Figure 3 relates to the relatively low level of the output gap in the highest demand periods. The output gap is close to or below 1 percent of the market capacity in the highest two demand categories, which corresponds to less than 250 MW. Even in a perfectly competitive market, the output gap cannot be expected to be zero because there will always be factors that cause marginal costs to fluctuate in a manner that is not reflected in the competitive benchmark, which will cause the estimated output gap to be greater than zero. Such factors include short-term fluctuations in fuel costs that are not perfectly accounted for by the fuel price adjustment of the reference price, changes in temperature or the technical characteristics of a resource that affects its efficiency, and changes in opportunity costs for resources that are subject to intertemporal output limitations.

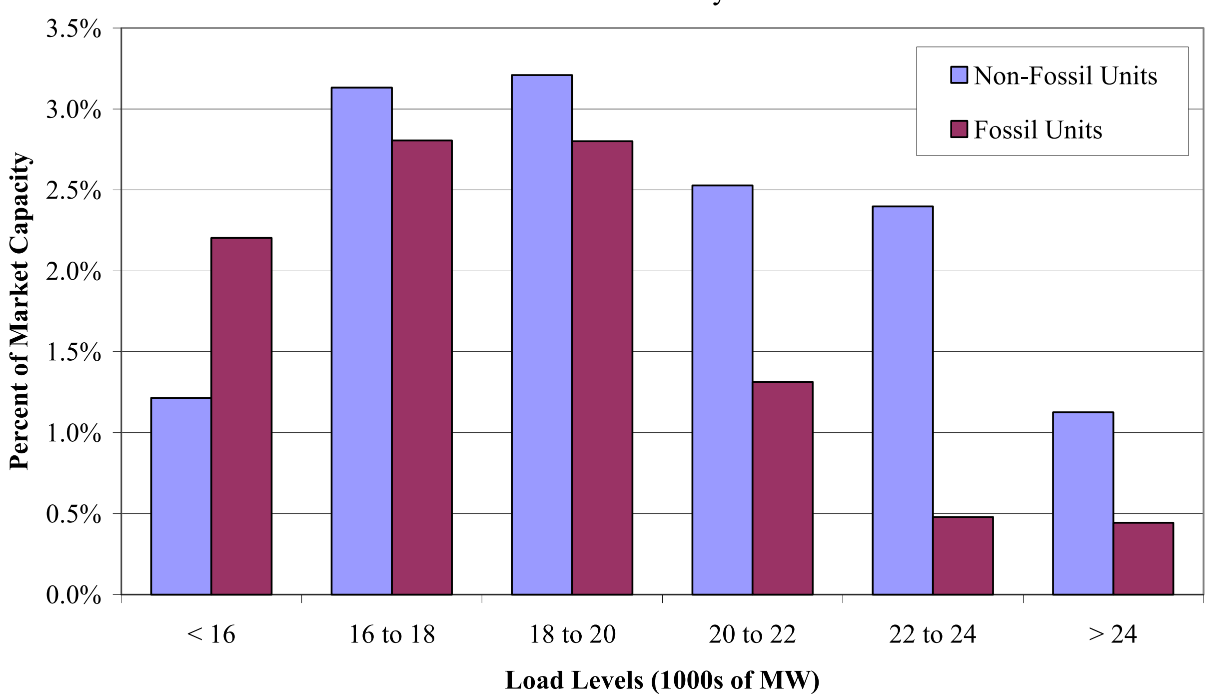
Indeed, the levels detected in this study at the highest demand levels are substantially lower than levels detected at lower demand levels when it is unlikely that significant incentives to withhold resources exist.

2. Unit Types

This section provides a comparison of the estimated output gap for fossil-fired units versus other types of generating units. With the exception of energy-limited fossil units and emergency capability at the top of the output ranges on some fossil-steam units, fossil units tend to exhibit more stable marginal costs than other units. This is particularly true with respect to hydro resources whose marginal costs may be largely comprised of the opportunity costs associated with shifting their production intertemporally (i.e., producing 1 MW this hour means that they cannot produce 1 MW in a future hour).

The stability of these costs allows the reference price methodology to produce a more reliable competitive benchmark and, therefore, reduces the probability of showing quantities in the output gap that do not reflect strategic economic withholding. Figure 4 shows the average

Figure 4
Average Output Gap by Type of Unit and Load Level
110% Reference Price Case: January to December 2001



Source: ISO New England Operations and Market Settlements Databases. Potomac Economics analysis.

output gap for fossil-fired and other units as a percent of the total market capacity corresponding to each fuel type. Almost two-thirds of the total market capacity in New England is fossil-fired capacity.

These results show that the output gap for all fuel types declines as the level of demand increases and that it is at its minimum in the highest-demand periods with the fossil units declining more rapidly. Because nuclear units rarely, if ever, contribute to the output gap, the quantities shown for other units generally represent hydroelectric resources. This result is not unexpected because a significant share of the hydroelectric resources exhibit marginal costs that are dominated by intertemporal opportunity costs for which reference prices are less accurate as a competitive benchmark for calculating the output gap. The balance of the analysis in this section relate to factors that may indicate the presence of market power or relate to incentives created by market rules in New England.

3. Three-Part Bidding

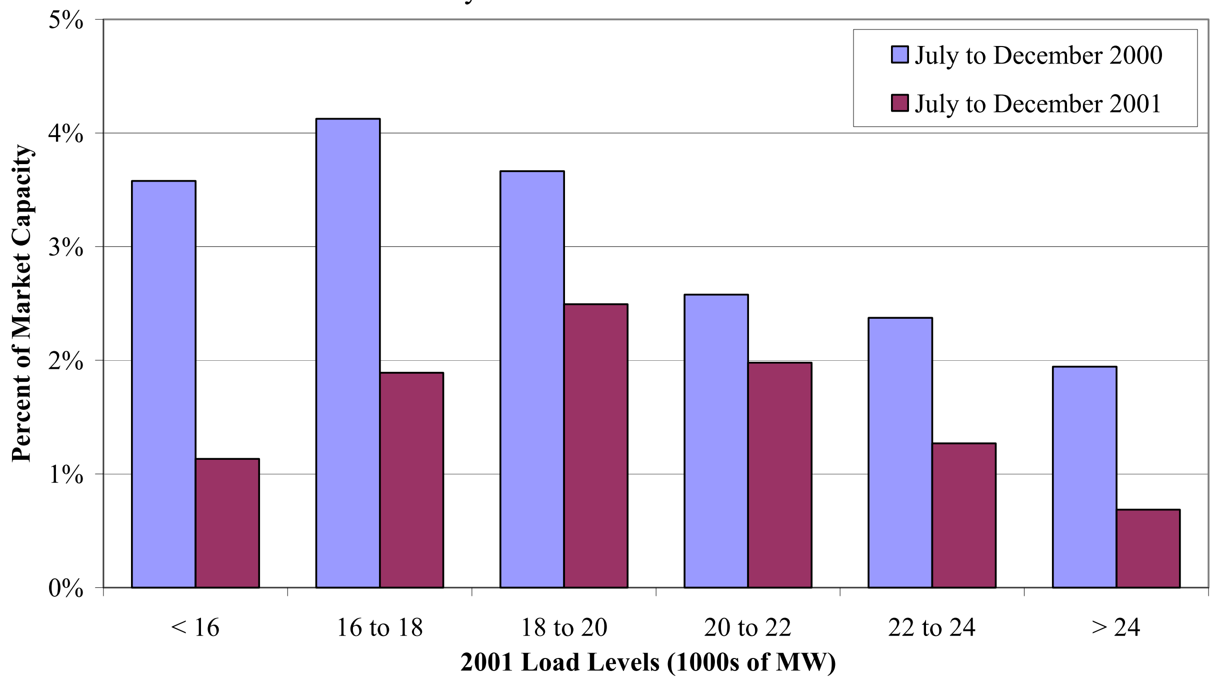
This section analyzes one of the key changes that occurred during 2001. When the market was first implemented in New England, suppliers submitted only an energy bid curve. This bid structure is referred to as a single-part bid because it includes only a bid for producing energy and does not include other bid components, such as the cost of starting-up the unit when it is off-line (start-up cost bid) or the fixed costs of producing at or above the minimum generation level (no-load bid).

Start-up cost bids and no-load bids play an important role in the decision to commit generation for the following day. In order to be economic, a unit must earn a sufficient margin over its incremental energy costs to recover its start-up and no-load costs. In a single-part bid structure, a unit that is clearly economic should continue to be bid at its marginal costs to ensure that its profits are maximized. However, owners of units that may not be economic may respond by raising their energy bids to ensure that they will cover their start-up costs if they are committed by the ISO.

As described above, the output gap statistics remove units that are truly uneconomic. However, units that are marginally economic that have raised their energy bids to account for their start-up and no-load costs and are not committed would be included in the output gap.

On July 1, the ISO's bidding rules were changed to allow for three-part bids that include an energy bid curve, a start-up cost bid, and a no-load bid. By including these latter two bid elements, suppliers would no longer have an incentive to account for these costs in their energy bid. To assess the impact that this rule change may have had on the output gap, I compared the post-July period in 2000 to the post-July period in 2001. This comparison is shown in Figure 5.

Figure 5
Average Output Gap Before and After Three-Part Bidding
July to December 2000 vs. 2001



Source: ISO New England Operations and Market Settlements Databases. Potomac Economics analysis.
Note: Hours for 2000 are allocated to the 2001 load level categories to match the distribution of the hours for 2001.

This analysis does not establish that the change in bidding behavior was caused by the implementation of three-part bidding on July 1, 2001. However, the results shown in this figure are consistent with expectations that the three-part bidding structure would reduce the incentive for suppliers to raise their bids above marginal cost.

Other elements of New England's market also affect generators' bidding incentives and, therefore, the estimated output gap. The most significant remaining element of the current market is the out-of-merit dispatch process.

4. Out-of-Merit Dispatch

Economists generally agree that generators in a competitive clearing-price market that lack market power should bid their marginal costs. Although the New England market sets hourly ECPs to settle spot market transactions, the Pricing Report showed that a substantial portion of the generation in New England is dispatched out-of-merit (i.e., dispatched even though the units' bid prices exceed the ECP). Most of these units are dispatched out-of-merit to manage transmission congestion and units in the congested areas tend to be dispatched out-of-merit much more frequently. Out-of-merit generators are paid their bid price rather than the ECP although some generators may be mitigated under Market Rule 17.

An alternative to the clearing price model for electricity markets is an "as-bid" market design where suppliers are paid their bid price when accepted rather than a market clearing price. In a perfectly competitive as-bid market, firms with no market power will rationally raise their bid to an equilibrium price level. For example, a coal unit with marginal costs of \$10/MWh and perfect foreknowledge that the competitive market price is \$30/MWh should be bid at \$30/MWh. Assuming no mistakes are made by the generators in forecasting the market price level, this will produce the same market result as bidding marginal costs in a clearing-price market since generators do not have to raise their bids in a clearing-price market to receive the market price.

In reality, the market price level will be uncertain because generators do not have full information regarding system conditions. The mistakes made in forecasting the market price generally cause as-bid markets to exhibit higher average prices and reduced economic efficiency relative to clearing-price markets.

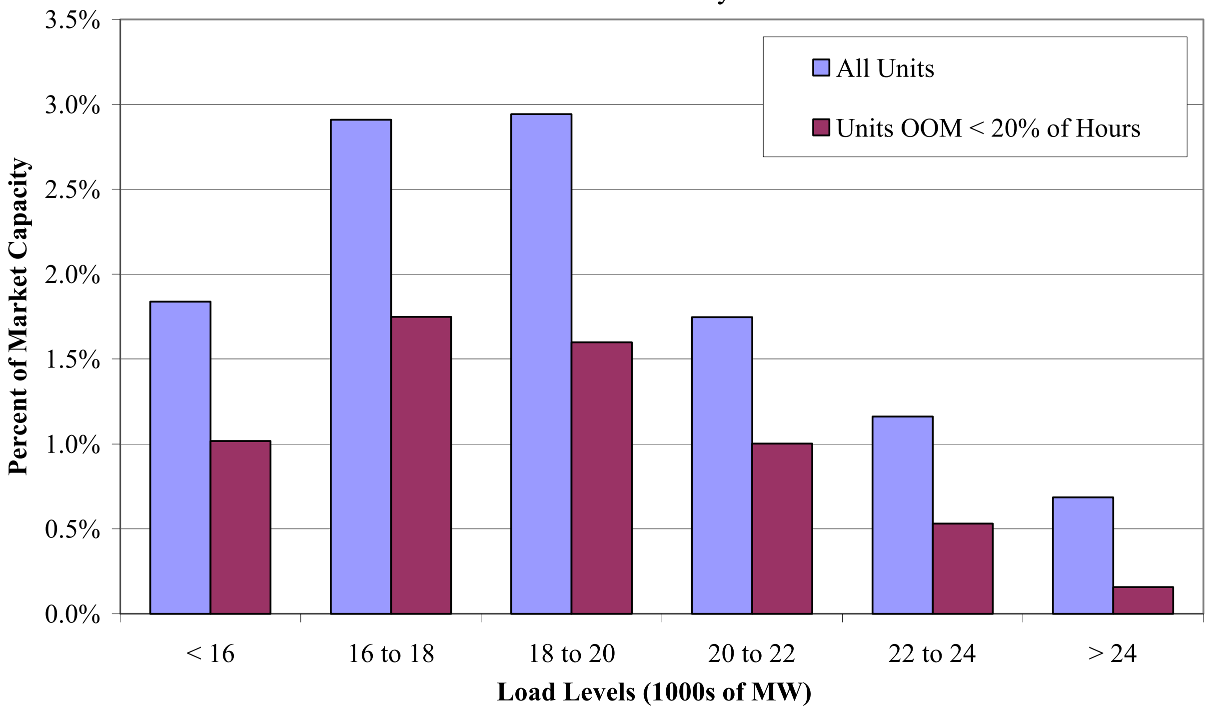
The fact that generators must raise their bids in an as-bid market is a function of the design of the market, not market power. These characteristics of the as-bid market design are described in the Blue Ribbon Panel Report produced for the California Power Exchange last year when some were considering restructuring its market from a clearing-price structure to an as-bid structure.[20]

Although bids by competitive suppliers in as-bid markets will rise above marginal costs, one cannot conclude that all increases in bid prices in as-bid markets are justified. For example, generators in congested areas in New England that are paid their bids and have increased their bids above marginal costs may be bidding competitively or may be exercising locational market power. ISO New England makes this distinction and imposes mitigation accordingly under Market Rule 17. Hence, this conduct is not analyzed in this report.

However, the out-of-merit bidding incentives can significantly affect the output gap estimates produced in this report. Generators that are frequently out-of-merit may raise their bid price in accordance with the as-bid incentives described above. When they are mistaken and do not run when the ECP is greater than their marginal costs, they will be added to the output gap.

Therefore, I conducted the output gap analysis for the subset of generators that generally run in-merit. The in-merit generators are defined as those that run in-merit at least 80 percent of the hours that they are dispatched (i.e., out-of-merit in less than 20 percent of the hours they run). Using this definition, approximately three quarters of the generators in New England are defined as in-merit. The results of this comparison are shown in Figure 6.

Figure 6
Effect of Out-of-Merit Dispatch on the Average Output Gap
110% Reference Price Case: January to December 2001



Source: ISO New England Operations and Market Settlements Databases. Potomac Economics analysis.

This figure compares the average output gap at each load level for the generators that are out-of-merit less than 20 percent of the hours versus all generating units. As in all of the figures, the percent of market capacity shown for each class of units includes only that class of capacity in the denominator. In other words, the output gap percentage shown for the in-merit units is computed by dividing the output gap from resources that are out-of-merit less than 20 percent of the hours they run by the total capability of all resources that are out-of-merit less than 20 percent of the hours they run.

The results in Figure 6 indicate that the output gap for in-merit units is substantially less under all load levels than the average for all units, which is consistent with the effects of out-of-merit dispatch on the bidding incentives described above. In addition, the relationship of the output gap to the load levels is consistent with the prior analyses showing that the output quantities are the smallest in the highest demand hours.

This analysis does not, however, establish that the out-of-merit bids have in all cases been competitive. If locational market power is present in the constrained areas in some of these hours, the increased output gap would also be consistent with an attempt economically to withhold by the participants in these areas. This report does not make this distinction, in part, because the as-bid incentives would be difficult if not impossible to differentiate from the market-power incentives. Additionally, Market Rule 17 allows the ISO to employ various bid screens to identify when an out-of-merit supplier faced limited competition in the constrained area and to mitigate its bid to address potential exercises of locational market power.

However, to the extent that the higher output gap for resources that are frequently out-of-merit is related to the as-bid incentives provided by the ISO's current congestion-management procedures, these incentives will be changed substantially when SMD is implemented. Under SMD, locational clearing prices are established in constrained areas that will give generators in these areas without market power the incentive to bid their marginal costs. In addition to changing the bidding incentives and sending more accurate price signals in constrained areas, SMD will make locational market power more transparent – easier to detect and to mitigate.

5. Participant Size

As discussed in the first section of this report, the two most important factors related to the ability and incentive of a participant to withhold resources to raise prices are the sensitivity of the prices to such withholding and the size of the participant. The size of the participant is important because it determines, in large part, whether the participant controls a sufficient quantity of resources to increase prices substantially and whether it would profit by doing so.

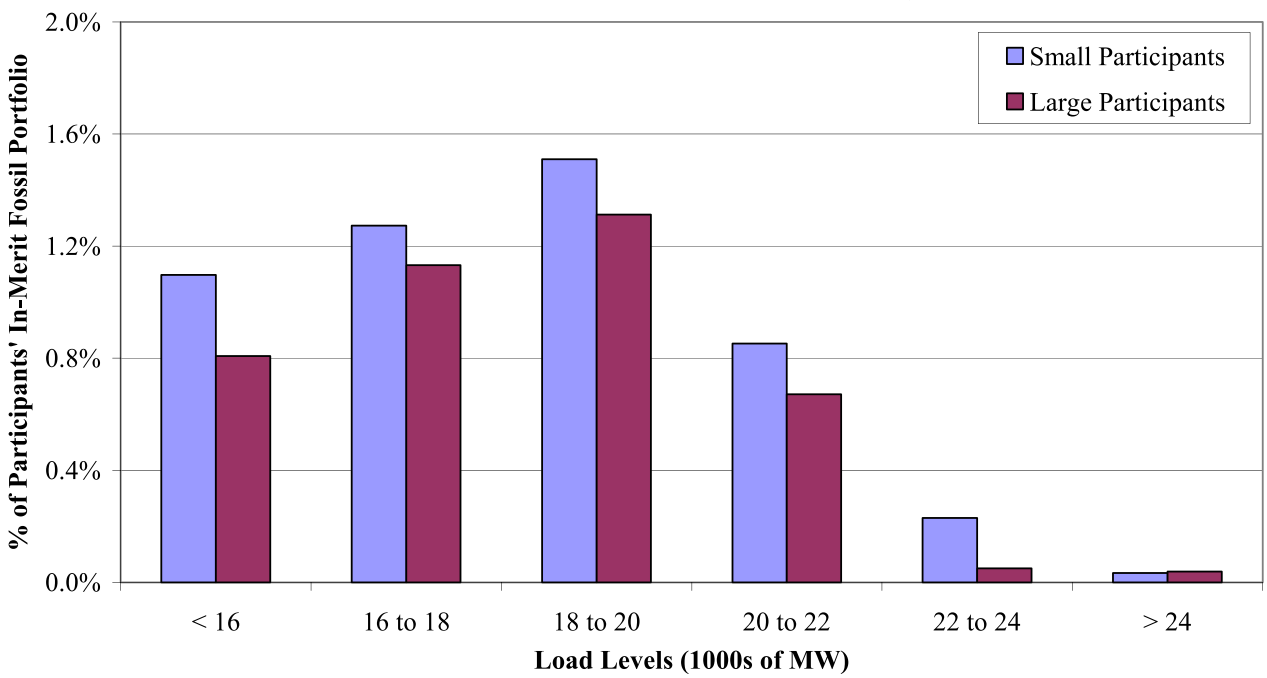
This section provides a comparative evaluation of the output gap for relatively large participants versus small participants. The econometric analysis presented below provides a superior means for determining whether the conduct of relatively large participants is consistent with workable competition because it controls for a number of significant factors, such as the fuel mix of generating portfolios.

As described above, for example, the output gap associated with hydroelectric resources is generally larger and more variable due to the fluctuation in marginal costs facing those units. In addition, it is important to account for the effects of out-of-merit dispatch on the output gap. Participants with larger portfolio shares that are frequently dispatched out-of-merit will tend to show larger average output gaps as shown in the prior section.

To account for these factors, the descriptive analysis presented in this section evaluates the average output gap for fossil-fired units (whose marginal cost characteristics are relatively stable) that are generally in-merit when they are dispatched (i.e., out-of-merit less than 20 percent of the hours they are dispatched). As described in the prior section, 75 percent of the generating units are generally in-merit when dispatched.

Figure 7 below shows the results of this analysis by size of participant. For purposes of this analysis, large participants are defined as those that own more than 1200 MW of available

Figure 7
Average Output Gap by Size of Participant During 2001
Fossil Units with Low Out-of-Merit Frequency



Source: ISO New England Operations and Market Settlements Databases. Potomac Economics analysis.

capacity. This will generally include the largest seven participants in the New England market. This level was selected because there was a natural division between the largest suppliers above this level and the other suppliers.

The figure shows three important results. First, the output gap as a percentage of the participants' portfolio of in-merit fossil units was slightly smaller for large participants than for smaller participants. This is important because small participants should serve as a benchmark for competitive behavior since they are much less likely to have market power, particularly when demand is not at super-peak levels. This figure indicates that the conduct of the largest participants does not exhibit higher levels of economic withholding than the smaller participants.

Second, the average output gap for both sizes of participants decreases considerably as the market approaches the highest demand periods. This result is important because the incentive to economically withhold resources, to the extent that it exists, should be much larger in the peak demand periods when prices are generally much more responsive to shifts in supply. This figure shows that the average output gap for the largest participants is close to zero in the two highest demand periods.

Lastly, the average output gap over all of the load levels is a very small share of the participants' portfolios – less than 1 percent on average. These results together are consistent with the hypothesis that the electricity markets in New England have been workably competitive.

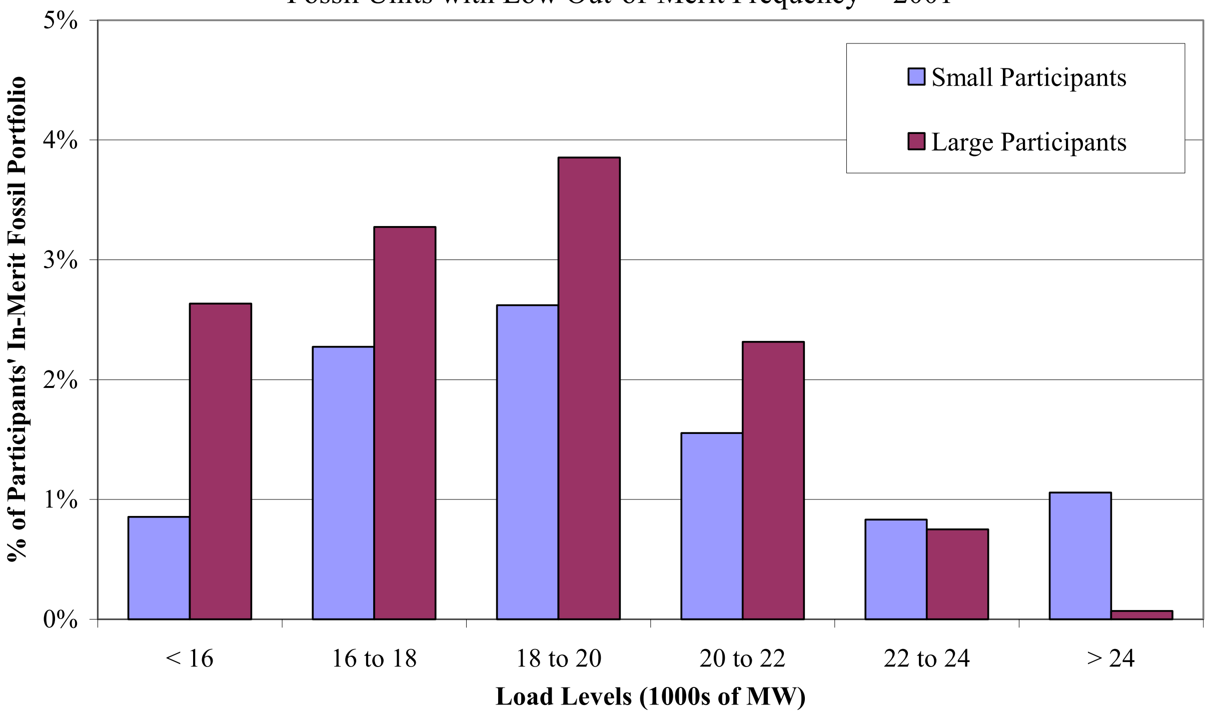
6. Alternative Competitive Benchmark

To determine whether these results are robust and not specific to the competitive benchmark used in the study, I have also estimated the output gap using variable production costs as the benchmark for each unit. Variable production costs have been used in a number of competitive analyses as a proxy for generators' marginal costs.

This assumption is reasonable for the majority of fossil-fired resources, however a significant portion of the market's capacity exhibits marginal costs that are substantially higher than variable production costs. For example, steam units often have an emergency output range that can only be achieved for a limited period of time by taking actions that may increase the variable O&M on the plant or increase its probability of incurring a forced outage. The additional exposure to expected losses from a forced outage and the increased O&M costs would contribute to the relatively high marginal costs of the emergency output.

Since all of these costs are incremental to the emergency output range and these ranges may only be the highest 1-3 percent of the unit's output range, the resultant marginal costs for the emergency output can be considerably higher than variable production costs. For this reason, the output gap estimated using the variable production costs is expected to be somewhat higher relative to the output gap based on reference prices. Figure 8 shows the output gap for large and small participants associated with fossil units that are generally run in-merit, which is comparable to the results shown in Figure 7 using reference prices.

Figure 8
Average Output Gap: 110% of Variable Production Cost Case
Fossil Units with Low Out-of-Merit Frequency - 2001



Source: ISO New England Operations and Market Settlements Databases. Potomac Economics Analysis.

As expected, the output gap estimated using the variable production costs is larger than the prior reference price results for all load levels and participant sizes. More importantly, these results confirm the prior results in that they show that the output gap declines substantially in the highest demand periods. In particular, the output gap is less than 1% for the largest participants in the highest two demand periods, and is considerably smaller than the estimated output gap during the same periods by smaller participants. Like the reference price results, these results are consistent with the hypothesis that the New England markets have been workably competitive and not subject to significant attempts to raise the energy clearing prices through economic withholding.

These results also show that in lower-load periods, larger participants have a larger output gap than smaller participants (which was not shown in the reference price results). Since this result only occurs in lower load periods when prices are much less sensitive to withholding, this result is not indicative of competitive concerns. It likely reflects differences in the generation owned by the large and small participants or differences in the bidding of particular output ranges. For example, large suppliers my own a larger share of the baseload fossil steam units with the emergency output ranges that I described above. These ranges would become part of the output gap when reference prices are replaced by variable production costs. Although a combined analysis including these other factors that may illuminate the output gap patterns is not possible within the analytic framework presented in this section, the econometric analysis presented in the following section does account for these factors.

D. Econometric Analysis of Economic Withholding

The descriptive analysis of the output gap from the preceding sections provides an overview of withholding behavior in the market and its relationship to key strategic and other variables that indicate whether the conduct detected is consistent with workable competition. However, the prior analyses are limited in revealing whether more complicated interrelationships exist among multiple factors.

The empirical hypotheses suggest that two factors should be positively correlated with the output gap if strategic economic withholding is present: high demand and large market participant size. The preliminary evidence in the descriptive analysis presented in the prior section suggests that the output gap is not positively correlated with either of these factors. However, if key market factors interact with one another, then the descriptive analysis may not reliably identify the specific impact of the key strategic variables.

A standard method of overcoming this problem is through econometric analysis. Econometrics applies statistical tools and procedures that can estimate the relationship of a number of independent variables to a dependent variable to be explained, the output gap in this case. The basic econometric tool is a regression analysis that estimates the influence of each of the independent variables on the dependent variable assuming a linear relationship between the independent variables and the dependent variable.

The relationship between the dependent variable and the independent variable is defined by the coefficient of the independent variable, b, in the equation: Y = a + b*X. The regression analysis would estimate the value of a and b in this example. In many cases, the standard error of the estimate is sufficiently large that one cannot determine with confidence that b is not equal to zero (i.e., that there is no relationship between the variables). Economists generally require that b is non-zero at the 95 percent confidence level before concluding that a statistically significant relationship exists between X and Y. [21]

The dependent variable in this econometric analysis is the output gap as a percentage of each participant's total capability in each hour. Therefore, the variables in this analysis correspond to the characteristic of the participant or the characteristic of the market in that hour.

1. Strategic Variables

The strategic variables included in the analysis focus on the size of participant and the demand level – the key factors derived from the supply function equilibrium presented in section II. To be more precise, the incentive to withhold occurs when supply is relatively inelastic (i.e., when the slope of the supply curve is steep). This only occurs at the highest demand levels. While there is no exact level of demand where the slope is clearly steep enough to motivate withholding, the top 1 percent of demand hours exhibit demands that are sufficiently high to cause the market to clear within or close to the inelastic portion of the supply curve. This frequency of demand is represented approximately by all hours when load is over 22,000 MW. Therefore, the first strategic variable is a dummy variable that identifies those hours when demand is higher than 22,000 MW in New England (*PEAK DEMAND*).

The second strategic variable identifies the largest participants in New England using these same definition used for the descriptive analysis. Thus, it is a dummy variable identifying those suppliers with more than 1200 MW of available capacity – generally the largest five to seven suppliers (*LARGE PARTICIPANT*). As described above, this level was selected because there was a natural division between the largest suppliers above this level and the other suppliers. However, even the largest suppliers are unlikely to have an incentive to withhold under moderate load levels when supply is elastic.

Therefore, the third strategic variable is a dummy variable that identifies the largest suppliers in the highest one percent of demand hours (*LARGE@PEAK*). This variable identifies the combination of factors that would maximize the incentive to economically withhold resources from the market and is, therefore, perhaps the most important strategic variable.

2. Non-Strategic Variables

Absent the market power explanation for the output gap, the other factors that would help explain the output gap generally relate to the measurement error associated with the competitive benchmark. In other words, the reference prices that serve as a proxy for each unit's marginal costs will not account for short-term fluctuations in the units' true marginal costs. Because data does not exist that would predict this measurement error, the regression analysis will not tend to explain a significant portion of the fluctuation in the output gap unless the output gap reflects an attempt to exercise market power, in which case the strategic variables should be highly significant.

Hydroelectric resources provide a good example of how the output gap can reflect measurement error related to detecting economic withholding. The opportunity costs associated with hydroelectric resources will fluctuate as market prices fluctuate. A hydro resource whose output must be managed on a weekly basis will incur higher opportunity costs in the warmest weeks during the summer when bidding in off-peak hours during these weeks. Therefore, reference prices based on seasonal data will tend to cause a larger share of these resources to be identified in the output gap when they raise their bids consistent with their higher opportunity costs. Hence, incorporating a variable in the regression analysis reflecting the portion of a participant's portfolio that is hydroelectric resources can explain some part of the output gap. In addition, including this type of non-strategic variable in the regression will hold this factor constant from participant to participant in estimating the relationship between the strategic variables and the output gap.

Failing to hold these types of factors constant could lead to spurious results associated with the strategic variables. If the some of the largest participants also happen to own a relatively large share of the hydroelectric resources, for example, the analysis may indicate that large participants exhibit larger output gaps due only to the hydroelectric factor. Therefore, by including these non-strategic variables in the regression analysis, the results regarding the strategic variables will be more reliable.

The non-strategic factors included in the regression analysis are summarized below:

- Composition of Market Participant Portfolio. To control for the effects of portfolio composition, the share of the portfolio that is comprised of base-load fossil plants, the share that is comprised fossil peaking plants, and the share that is comprised of hydroelectric plants are all put into the analysis.[22] The average age of the plants in each portfolio is also included in the analysis.

- Seasonal and Time-of-Day Factors. The regression also includes a variable indicating the season, allowing control for changes in output gap arising from seasonal factors. It also includes a variable indicating whether or not it is a work-hour or nighttime/weekend/holiday -- (work hours are between 6AM and 10PM weekdays and non-holidays).

- Propensity of Portfolio to be Out-of-Merit. In order to adjust for the effect of out-of-merit dispatch on participants' bidding behavior, the share of the market participant's portfolio that is dispatched out-of-merit more than 20% of the time when they are running is included as a variable. Frequent out-of-merit dispatch will give participants the incentive to raise their bids above marginal cost, increasing the likelihood that they appear in the output gap.

- Fuel Prices. The output gap is measured based on estimates of the economic level of output versus actual output. The economic level of output changes as fuel prices change. While estimates of reference prices are adjusted to take into account changes in fuel prices, this adjustment is inevitably imperfect and can affect the estimated output gap.

In addition to these non-strategic variables, the model includes the output gap value for the preceding period. This lagged variable is included to account for the serial correlation in the output gap and may reflect the fact that bids are likely not independently formulated in each period. The factors influencing bidding decisions in one period are likely to influence the bids in following periods. If the model does not account for this, it will exhibit substantial autocorrelation. Without the lagged variable to account for the serial correlation in the output gap, the model could estimate spurious relationships with other independent variables that are also serially correlated. The lagged variable does substantially address the autocorrelation in the model.

Appendix C describes the regression analysis in more detail and presents the full results of the analysis. The Appendix also includes two other cases that were estimated to test the robustness of the results of the base model. The first is a case that includes only the hour beginning 3 p.m. for each day in the analysis. This case tests whether conduct occurring in off-peak hours influences the estimated relationship between the output gap and the strategic variables.

The second case tests whether the same relationship holds when only fossil units are included in the analysis. The report includes this case because the marginal costs of fossil units are generally more stable and hence less subject to measurement error than other types of units. Therefore, assessing whether the same relationships hold for the fossil units tests the robustness of the results of the base model.

3. Regression Results

The full results of the regression analysis are presented in Appendix C. Table C1 shows the results of the three cases – the base model, 3 p.m. hours only, and fossil units only. As described above, the key regression results on which to focus for the purposes of the empirical hypotheses are the estimated coefficients of the three strategic variables: (1) the variable *PEAK DEMAND* -- an indicator variable associated with the top 1% of demand hours; (2) the variable *LARGE PARTICIPANT* -- an indicator variable representing large market participants (with portfolios greater than 1200 MW); and (3) the interactive term *LARGE@PEAK* that indicates both high-demand hours and large participant.

The *PEAK DEMAND* variable shows a statistically significant negative relationship to the output gap at the 95 percent confidence level in all three cases, with estimates ranging from -0.003 to -0.014. Therefore, these results indicate that, *ceteris paribus*, the portion of a participants' portfolio included in the output gap is generally approximately 1 percentage point *lower* in the peak demand hours than in other hours. This result is consistent with the descriptive analysis presented in prior sections and with the hypothesis that the New England markets have been workably competitive.

The *LARGE PARTICIPANT* variable is statistically insignificant in the base model and the 3 p.m. model. However, it is shows a statistically significant negative relationship to the output gap in the fossil units case with an coefficient of -0.0015. This result means that the portion of a large participants' portfolio included in the output gap is generally 0.15 of a percentage point less than the comparable portfolio portion for smaller participants. Again, this result is consistent with the descriptive analysis presented in prior sections and with the hypothesis that the New England markets have been workably competitive.

The regression model estimate for the *LARGE@PEAK* variable is statistically insignificant in all three cases.[23] This is consistent with the hypothesis that the New England markets have been workably competitive because attempts to exercise market power in these periods would create a statistically significant positive relationship.

Hence these results support the conclusion that the markets have been competitive and not subject to systematic attempts to economically withhold resources to raise market prices. However, this addresses only one form of withholding. The next section analyzes physical withholding.

IV. Physical Withholding

A. Introduction

This section analyzes generator availability to assess whether strategic physical withholding was a significant concern during 2001. A resource may be physically withdrawn from the market by derating the generating unit, i.e., lowering the unit's high operating limit ("HOL"). There are generally two categories of generator deratings – generator outages where the HOL is generally reduced to zero, and other deratings where the HOL is set at a positive value below the unit's maximum capability.

Because unit outages and other deratings would both occur in perfectly competitive electricity markets, the analytic objective of this section is to determine the extent to which the empirical evidence regarding these deratings is consistent with workable competition. Like the output gap analysis in the prior section, this section analyzes deratings relative to market conditions and participant characteristics to differentiate strategic withholding from naturally occurring outages.

B. Outage and Other Deratings Data

1. Planned and Forced Generator Outages

The first means of physically removing supply from the market is by declaring an outage, of which there are two primary classes: planned outages and forced outages. Planned outages are scheduled to perform routine and other maintenance that is not urgent. Under FERC policy, these outages should be coordinated by the ISO. Hence, planned outages are removed from the outages used for the withholding analysis since it is unlikely that generators seeking to strategically withhold resources would do so with planned outages since these outages are scheduled by the ISO.

Further, planned outages are generally scheduled in off-peak periods. Therefore, including them in the analysis would bias the results toward showing higher generator availability in peak periods and, in doing so, obscure the identification of strategic physical withholding.

The second class of outages is forced outages that are not scheduled by the ISO, although the ISO is notified by the generator. Although forced outages are a normal occurrence in electricity markets, they are also one means for strategically withholding resources from the market. To focus the analysis on the forced outages that are most likely to be strategic, I have divided the outages into short and long-term outages, defining long-term outages as those with a duration longer than 7 days.

Long-term forced outages include more than half of the forced outages and are less likely to be used to strategically withhold resources due to the costs of foregone market revenues in hours when the supplier does not have significant market power. Further, excluding long-term outages is very important for this analysis because longer-term outages are naturally weighted more heavily than shorter-term outages even though they are less likely to be strategic.

For example, a one month outage will receive 30 times more weight than a 1 day outage, and 180 times more weight than a 4 hour outage even though the short-term outage occurring under peak demand conditions may be the most likely to be strategic. Hence, the analysis in this report includes only short-term forced outages to focus it on the conduct that is most likely to constitute strategic physical withholding. Assuming for the moment that strategic withholding occurred during the study period, focusing the analysis on forced outages of 7 days or less will minimize the possibility that the conduct will be masked by long-term outages.

In addition, the short-term forced outages analyzed in this study have been subdivided into those with a duration of one day or less versus multiple day outages. Following much the same logic as on long-term outages, intraday outages would be the most profitable means of withholding since they would not incur lost revenues in adjacent periods by not being available. Due to the weighting in the averages, the impacts of the intraday outages may be masked by the multiple day outages if not separately identified.

The recent studies of market power in the California market employ more restrictive assumptions regarding the types of forced outages that may constitute strategic physical withholding. The least restrictive case used by Joskow and Kahn counted forced outages as strategic only if they did not also occur on the day before the day in question.[24] To ensure that the inclusion of long-term outages up to 7 days does not mask strategic conduct associated with forced outages occurring intraday, I have subdivided the 7 day outages into those of one day or less versus those of longer than one day. This allows the analysis in this report to separately assess the patterns as they relate to each type of outage.

2. Other Deratings

The second category of conduct that can be used to physically withhold resources from the market can be referred to as "other deratings". This includes all cases where the high operating limit ("HOL") is reduced below the maximum capability and there is no log entry to indicate that it corresponds to a forced or planned outage. These types of deratings occur each day for virtually every generating unit in the market simply because units are generally not always capable of achieving their maximum output level. For example, ambient temperature conditions can cause fossil units to reduce their HOLs from the maximum output level that can be achieved under ideal temperatures.

The one adjustment made to the derating data is made to account for resources that are congested down. Generating units whose output must be limited because they are located in export-constrained regions, such as Maine, will generally be derated by the ISO with a log entry identifying the system condition. These deratings are removed from the data since they are done by the ISO rather than the participant.

Finally, the maximum capability for each unit that is used to compute the deratings and outage values is estimated by taking the minimum of: the seasonal claimed capability, the highest HOL set during the season, or the highest quantity of energy bid during the season. This approach is preferred to using only the seasonal claimed capability because some units have claimed capabilities that exceed the true maximum capability of the unit, and thus avoids overestimating the true level of deratings.

Using this methodology for calculating the deratings and short-term forced outages, this data is then used for the physical withholding analysis below. The result of the descriptive analysis of the deratings and outages is presented in the next section while the econometric analysis of this data is presented in the following section.

C. Results of the Physical Withholding Analysis

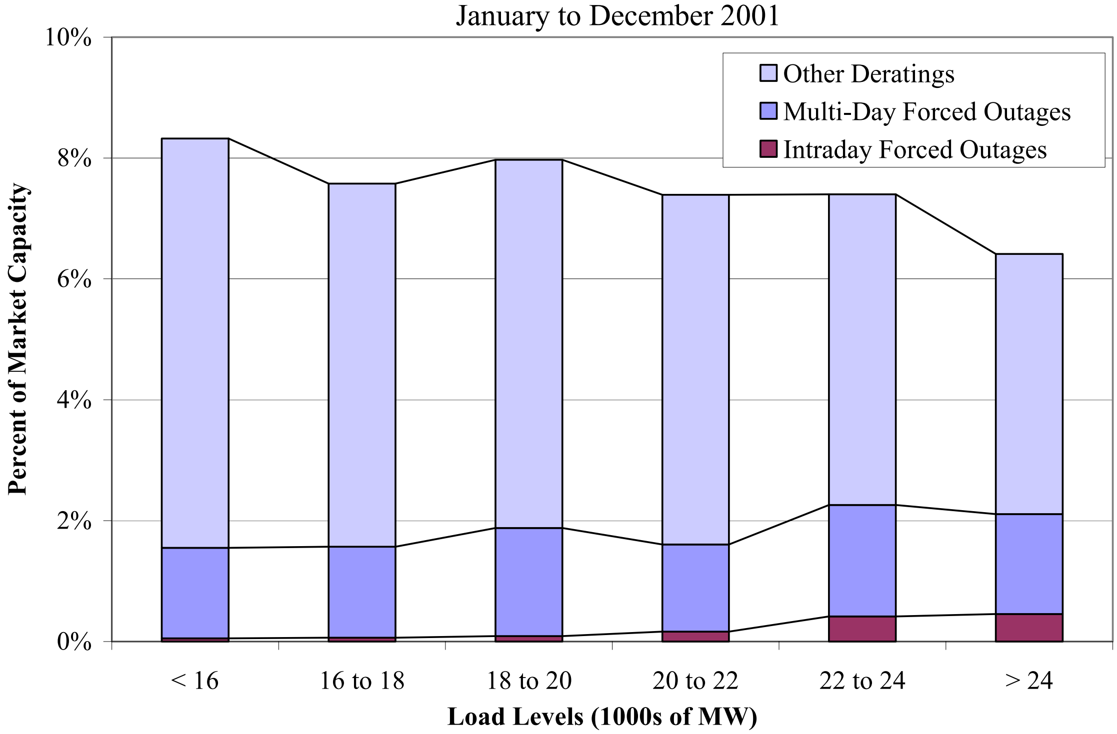
Having established the outage and other derating values for each unit on an hourly basis, this data is used to identify the extent to which strategic physical withholding may be a significant concern in the New England market. This analysis is very similar to the output gap analysis presented above in that it evaluates derating quantities relative to market conditions and participant characteristics.

As described above, strategic withholding should occur only when certain factors are present that create the ability and incentive for one or more suppliers to exercise market power. The analysis of the deratings relative to these factors are presented in the following subsections.

1. Demand Levels

The first analysis of physical withholding assesses the relationship between the deratings and demand levels. As discussed above, the market's vulnerability to exercises of market power should be substantially higher when prices are likely to be more responsive to withholding. The same load levels are used to compute the average amounts of outages and deratings as were used in the output gap analysis in the prior section. Figure 9 shows the average quantity of forced outages and other deratings by load level.

Figure 9
Forced Outages and Other Deratings by Load Level
January to December 2001



Source: ISO New England Operations and Market Settlements Databases. Potomac Economics Analysis.

The results shown in this figure are generally consistent with workable competition in that one would expect suppliers to maximize the availability of their generation in the peak hours when prices are likely to be the highest. This is accomplished by minimizing the deratings over which participants have some control. For example, a participant might make output available from a generating resource that it would not want to produce on a regular basis, but which would be profitable under peak conditions. One clear example of this are energy limited resources, such as hydroelectric resources, that will seek to shift their output to the highest value hours.

Alternatively, legitimate forced outages should be random and should, therefore, not exhibit the same negative correlation to load levels as the other deratings. In fact, one may expect a positive correlation between demand levels and forced outages as units are run at higher levels for sustained periods and other relatively high-cost peaking resources are dispatched that seldom run. In both cases, the dispatch of resources that are not generally utilized at lower load levels may result in a higher frequency of forced outages during peak hours. This is consistent with the total level of forced outages shown in figure 9. However, this also makes it more difficult to differentiate between legitimate forced outages and strategic withholding, thereby emphasizing the importance of the ISO's physical audit program.

Slightly higher levels of forced outages have occurred during some of the relatively high demand hours (load levels of 22,000 MW to 24,000 MW), some of which could correspond to strategic physical withholding. However, the total forced outages are reduced when the load level was greater than 24,000 MW.

With respect to the intraday outages shown in the figure, these outages are the one class of forced outage that is significantly higher in the peak demand periods than in the off-peak periods. These are also the class of forced outage that would be the most attractive strategy to a supplier attempting to raise prices since the withholding can be focused on the hours that are most likely to be susceptible to an exercise of market power. Hence, this issue warrants further analysis, which is provided in the following sections.

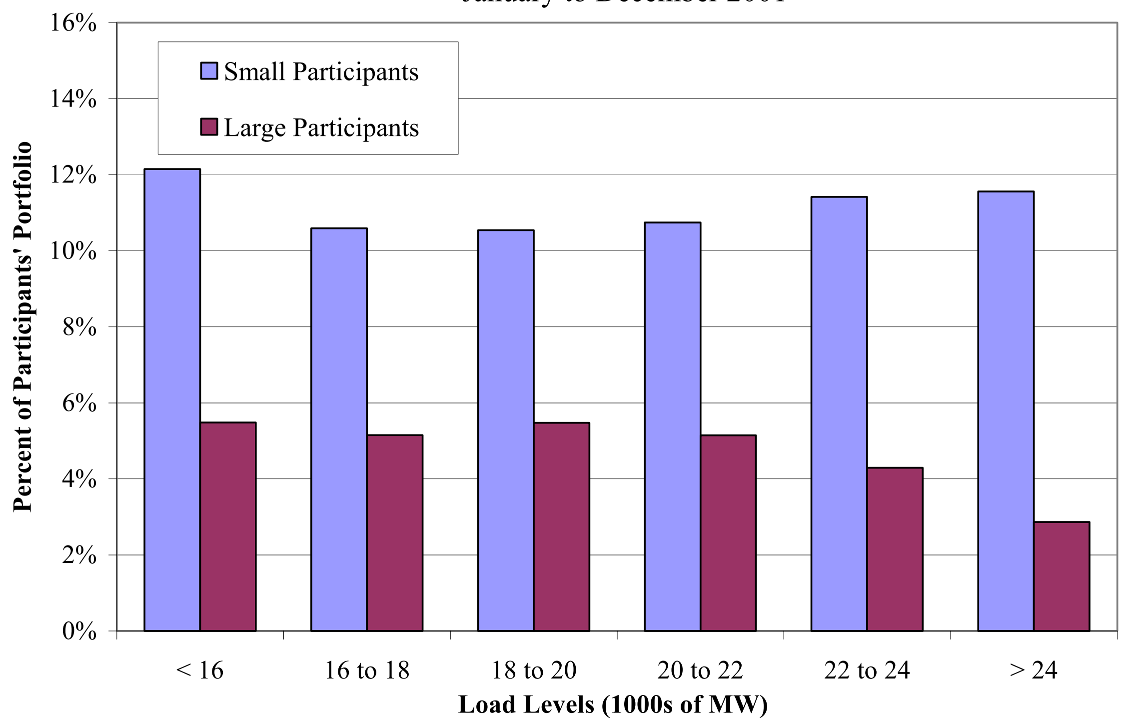
2. Participant Size

This section of the report analyzes the relationship of participant size to the outage and derating levels. As discussed above, the total capacity owned by a participant is one of the key determinants of whether the supplier has the ability and incentive to raise market prices by strategically withholding.

The analysis in this section builds upon the analysis of demand levels above in that each of the factors analyzed below is shown relative to the demand levels established in the prior section. However, the analysis in this section differs in that it shows the deratings and outages as a percentage of the participants' available portfolio while the analysis in the prior section showed these values as a percentage of the total market capability. The only difference in these measures is the weighting that is applied to participants of different sizes (i.e., all suppliers receive the same weight in this analysis where as large suppliers were weighted in proportion to their portfolio size in the prior section).

The definition of large participants employed in this section is the same as the definition employed in the output gap section – participants with available capability greater than 1200 MW are categorized as large participants. This generally identifies the largest seven suppliers within New England. The first analysis, shown in Figure 10 below, evaluates the differences in other deratings between large and small participants.

Figure 10
Deratings Other Than Forced Outages by Participant Size
January to December 2001



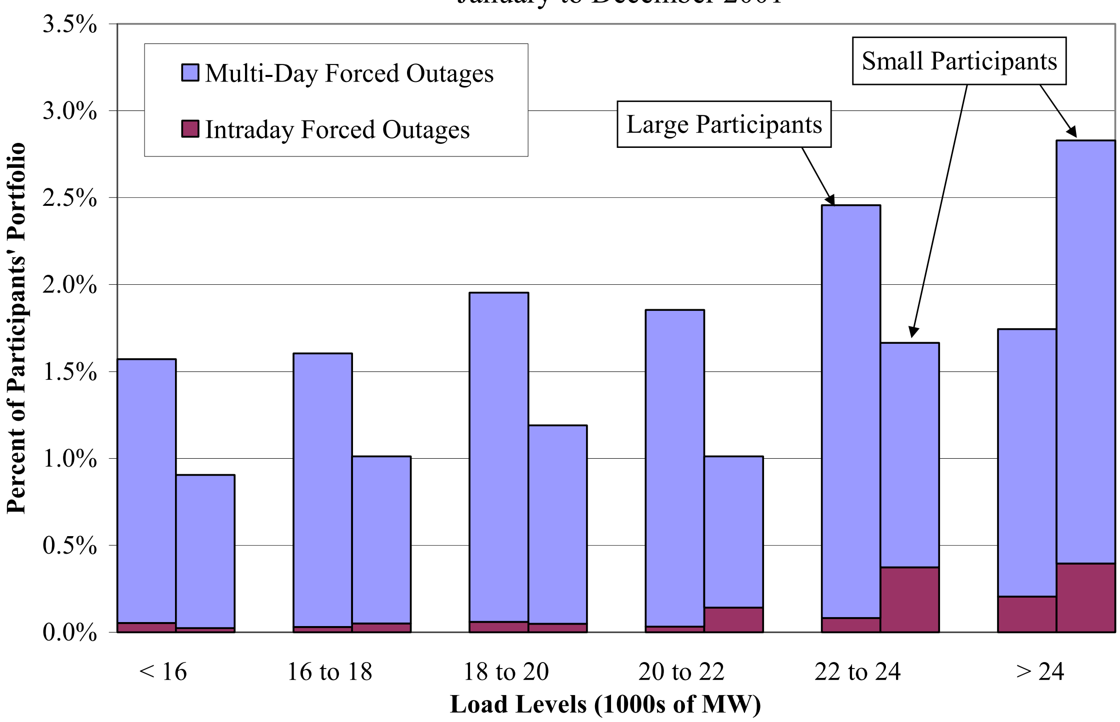
Source: ISO New England Operations and Market Settlements Databases. Potomac Economics Analysis

As the figure shows, large participants at all load levels exhibit substantially lower non-outage related deratings than small participants. This could reflect efficiencies that are achieved in managing a relatively large portfolio of generating assets, or differences in the composition of the portfolios owned by large suppliers. Some of these factors are best accounted for in the econometric analysis presented below.

Additionally, Figure 10 shows that the other deratings are lower in peak demand hours for the large participants. This supports the workably competitive hypothesis that suppliers lacking market power will seek to maximize their availability in peak hours when market prices are the highest. I applied the same analysis to the forced outage data to test the same hypothesis.

Figure 11 focuses on the differences in forced outages in 2001 between large and small participants. As described above, the outages in this analysis are divided between intraday outages versus multi-day outages.

Figure 11
Forced Outages by Participant Size
January to December 2001



Source: ISO New England Operations and Market Settlements Databases. Potomac Economics Analysis.

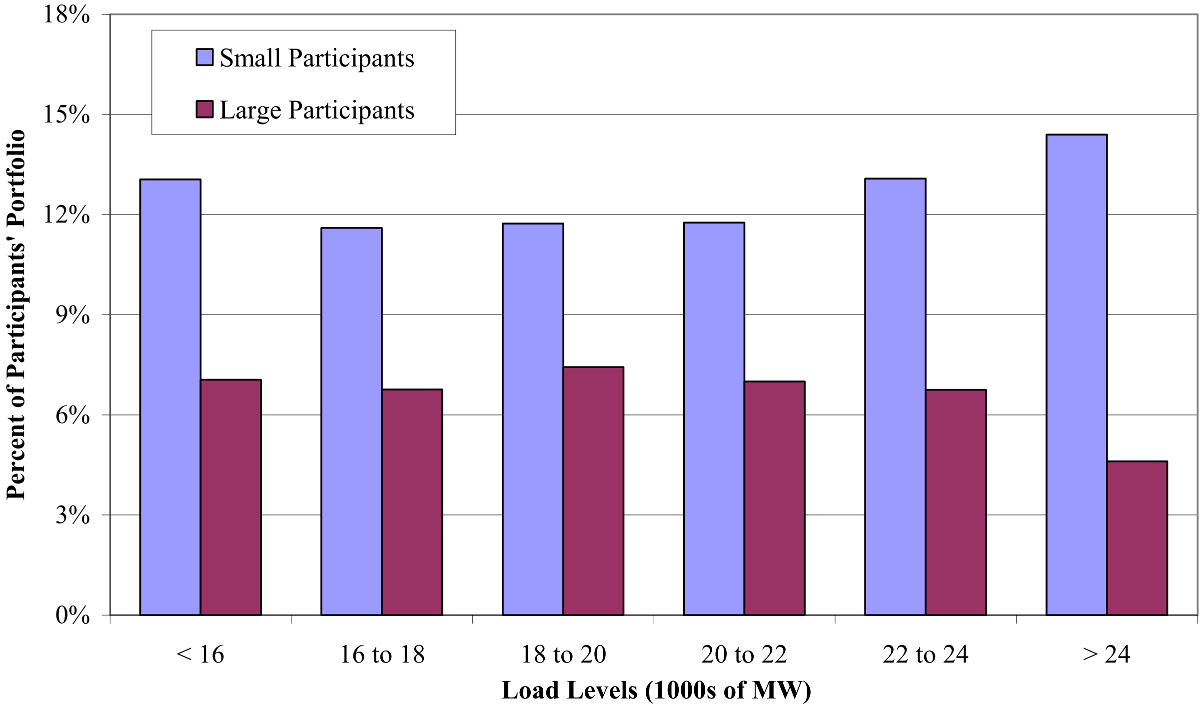
This figure shows that the forced outage rates for the large suppliers are higher on average for all load levels, with the exception of the super-peak demand. This is not indicative of strategic physical withholding for three reasons. First, the fact that comparable differences occur in the lowest demand periods as in the higher demand periods indicates that this difference is likely related to other factors, such as differences in technology mix or outage reporting practices. Second, in the periods where the incentive should be the highest to withhold – the super-peak period – the forced outages by large participants were substantially less than by small participants. Third, these differences are substantially smaller than the differences in other deratings between large and small participants, causing the total deratings values for large participants to be lower than for small participants.

However, the increase in intraday outages for both large and small participants that occurs in the highest demand periods does justify further investigation since it is correlated with those conditions when withholding would likely be most profitable. The econometric analysis presented later in this section analyzes the intraday outages in more detail to determine whether these outages were likely strategic or whether they are consistent with competitive expectations. Regardless of the results, however, one cannot conclude with certainty that none of the forced outages was a strategic attempt to raise prices. Therefore, the ISO's monitoring and audit program to detect and identify strategic physical withholding should remain a relatively high priority.

Figure 12 combines the other deratings and forced outages to provide a more complete picture of the total deratings by large and small participants. These results show that the total deratings by large participants is consistently lower than the total deratings by smaller participants. Again, this may indicate the presence of operating efficiencies in owning a larger portfolio of assets or reflect differences in portfolio mix of generating technologies. The figure also shows that while total deratings by small participants rise slightly in the highest demand hours, the total deratings by large participants do not.

This combined analysis is meaningful because it indicates whether the forced outages and the deratings together support a conclusion that resources may have been strategically withheld. This is the most meaningful result because large suppliers attempting to raise prices by claiming forced outages strategically would employ a consistent strategy with their deratings (i.e., increasing the deratings). Figure 12 shows that the deratings by large participants more than offset the increase in forced outages and support the conclusion that the forced outages were generally non-strategic.

Figure 12
Total Deratings by Participant Size
January to December 2001



Source: ISO New England Operations and Market Settlements Databases. Potomac Economics Analysis.

The econometric analysis presented in the following section will better assess these relationships by controlling for the effects of other key factors.

D. Econometric Analysis of Outages and Other Deratings

The econometric analysis of outages and other deratings presented in this section is similar to the econometric analysis of the output gap. In particular, we use linear regression models to test whether deratings are related to the strategic variables, high demand and large market participant size, that would indicate that the deratings may be an attempt to exercise market power by physically withholding supply.

1. Dependent and Independent Variables

The dependent variables in these models include the various types of outage and deratings as a percentage of each participant's total capability in each hour. The four models tested are:

- Base Model – the dependent variable equals the total deratings, which equals the forced outages less than 7 days plus other deratings, as a percent of the participants' portfolios.
- Other Deratings – the dependent variable includes only the other deratings as a percent of the participants' portfolios.
- Forced Outages – the dependent variable includes all of the short-term forced outages as a percent of the participants' portfolios.
- Intraday outages– the dependent variable includes only the Intraday outages as a percent of the participants' portfolios.

This analysis includes substantially the same non-strategic variables in the regression analysis for physical withholding as in the output gap regressions because these variables may also affect deratings, although possibly in different ways and for different reasons. Like for the output gap analysis, the analysis includes market participant portfolio characteristics, including the composition of the portfolios and average age because these factors likely explain some of the differences in derating quantities. We also include seasonal variations and peak day (non-weekend and holiday) indicators. However, the analysis does not control for fuel prices – while they are likely to influence the output gap, they have no analytic connection to deratings.

Like the output gap analysis, the model includes the lagged value of the dependent variable in the preceding period. This lagged variable is included to account for the serial correlation in the deratings data, which is generally more severe on an hourly basis than the serial correlation in the output gap. To more fully address the autocorrelation that this causes and exclude deratings that are taken by hydro units in off-peak hours, all of the models are run including only the 3 p.m. hour of each day. This focuses the analysis on the changes in deratings occurring in the daily peak hours to improve the reliability of the regression results. Appendix C describes the regression analysis in more detail and presents the full results of the analysis.

2. Regression Results

Table C2 in Appendix C shows the results of the first three cases – the base model, the other deratings case, and the forced outage case. As described above, the key regression results to focus on for the purposes of the empirical hypotheses are the estimated coefficients of the three strategic variables: (1) the variable *PEAK DEMAND* -- an indicator variable associated with the top 1% of demand hours; (2) the variable *LARGE PARTICIPANT* – an indicator variable representing large market participants (with portfolios greater than 1200 MW); and (3) the interactive term *LARGE@PEAK* that indicates both high demand hours and large participant.

The *PEAK DEMAND* variable is statistically insignificant in all three cases at the 95 percent confidence level. This result is consistent with the descriptive results shown above and cancels each other in the base model. The prior sections provide possible explanations for the decrease in deratings and increase in forced outages that occur in the super-peak demand periods.

In particular, the increase in forced outages in these periods may by attributable to the fact that older, more costly, and less reliable units are dispatched in these hours, or may reflect strategic withholding. However, it is important to recognize that these statistical results are not reliable in determining that these relationships actually exist since they cannot be established at the 95 percent confidence level. Further, it is reassuring that the results of the base model show no significance for the *PEAK DEMAND* variable since one would expect that strategic forced outages would be complemented by strategic deratings rather than being offset by decreased deratings.

The *LARGE PARTICIPANT* variable is negative and statistically significant in the base model and the other deratings model with coefficient estimates of -0.02 and -0.023. This result means that the portion of a large participant's portfolio that is derated is generally 2 percentage points less than the comparable portfolio portion for smaller participants. This result is consistent with the descriptive analysis presented in prior sections and with the hypothesis that the New England markets have been workably competitive.

However the *LARGE PARTICIPANT* variable is positive and statistically significant in the forced outage case with a coefficient of 0.0033. Hence, the portion of a large participants' portfolio that is forced-out is generally one-third of a percentage point more than the comparable portfolio portion for smaller participants. In addition to this estimate being relatively small, this result is inconsistent with the deratings results, and is inconsistent with the next results analyzing the conduct of large participants under the super-peak demand conditions.

The regression results for the *LARGE@PEAK* variable is statistically insignificant in all three cases. This is consistent with the hypothesis that the New England markets have been workably competitive because attempts to exercise market power in these periods would result in a significant positive relationship.

Taken together, these results support the conclusion that the markets have been competitive and not subject to systematic attempts to physically withhold resources to raise market prices. However, the report includes one last case that focuses on the intra-day forced outages. These outages would be the least costly to employ strategically as a means to raise prices and are, therefore, of particular interest.

The results of this case are shown in C3 in Appendix C. It shows that none of the variables is significant with the exception of the *PEAK DEMAND* variable.[25] In addition, the model as a whole explains almost none of the forced outage activity. This is expected since these outages should be random.

The *PEAK DEMAND* variable is positive and significant in this case (as opposed to the total forced outage case). The results indicate that, *ceteris paribus*, the portion of a participants' portfolio on intraday outage is generally 0.3 percentage points higher in the peak demand hours than in other hours. Again, there are strategic and non-strategic explanations for this result and the only means available to determine whether some of these outages were strategic is through an auditing of the units. This program is valuable in that it allows the ISO to directly monitor for this form of withholding and significantly increases the deterrent to engaging in this conduct.

However, the fact that the other two strategic variables show no statistical significance is positive evidence suggesting that the results for the peak demand variable are not indicative of strategic withholding by the participants.

V. Analysis of Highest-Priced Hours During Summer 2001

This section provides a detailed analysis of the highest-priced hours during the Summer 2001. For purposes of this analysis, I selected all hours priced at $200 per MWh or above. The hours selected under this criteria represent the highest 0.2 percent of the hours during the summer.

As discussed above, the nature of the supply and demand in the current wholesale spot electricity markets is such that prices will be far more sensitive under super-peak conditions to withholding or market rules that do not facilitate full utilization of the system's resources. Since prices in these hours can be many times larger than the average price, the costs associated with unjustified price increases can be large even when the periods exhibiting these prices are relatively infrequent and short-lived.

As the Pricing Report indicated as well, the ability of the market to establish efficient prices under peak conditions when the market is tightest is critical to both consumers and suppliers. Thus, the focus of the analysis in this section on these hours is warranted. Table 3 provides a summary of the hours evaluated in this section.

In most of these hours, the ECP is set at $1,000 by external transactions. Under the rules that prevailed last summer, all dispatchable external transactions that are accepted by the ISO set a floor on the energy price in New England.[26] In every hour showing a $1,000 price in the table above, an external transaction was setting the price. As described in the Pricing Report, this rule is justified only to the extent that the external transaction is the most economic means of meeting the ISO's energy and reserve requirements. The analysis in this section evaluates whether this was the case for each of the high-priced hours shown in Table 3. The reforms filed by the ISO in response to the Pricing Report are intended to ensure that out-of-merit external transactions are accepted when they are the marginal economic resource and set prices accordingly in the future. These reforms are described in more detail below.

This section will also evaluate whether any of the high prices that occurred during the Summer 2001 would have been prevented by the pricing reforms filed by the ISO. This section also assesses the extent to which the transaction scheduling provisions in adjacent markets may also have contributed to the high prices shown in Table 3.

Table 3
New England Market Summary for High Priced Hours
Hours with ECP > $200 During Summer 2001

Date and Time	Load	Energy Clearing Price	Externals Accepted @ $1000	Reserves Shortfall
7/23/01 - 6 PM	21,447	$1,000	288	0
7/23/01 - 7 PM	20,742	$1,000	288	0
7/24/01 - 10 AM	21,723	$226	0	0
7/24/01 - 1 PM	23,554	$1,000	321	-53
7/24/01 - 2 PM	23,656	$1,000	334	-121
7/24/01 - 3 PM	23,653	$1,000	352	-38
7/25/01 - 12 AM	23,657	$1,000	352	0
7/25/01 - 1 PM	23,957	$1,000	352	-12
7/25/01 - 2 PM	24,085	$1,000	352	-301
7/25/01 - 3 PM	24,189	$1,000	352	-391
7/25/01 - 4 PM	23,998	$1,000	352	-115
7/25/01 - 5 PM	23,466	$1,000	352	-289
7/25/01 - 6 PM	22,879	$1,000	352	0
7/25/01 - 7 PM	22,329	$1,000	352	0
8/9/01 - 12 AM	24,725	$1,000	352	0
8/9/01 - 1 PM	24,951	$1,000	33	-249
8/9/01 - 3 PM	24,918	$243	0	-1,120
8/9/01 - 4 PM	24,735	$217	0	-847

Source: ISO-NE Operations and Market Data. Potomac Economics analysis.

A. Energy and Reserve Requirements

This section provides the analysis of the supply and demand conditions in both the energy and reserve markets during the high-priced hours. Examining these conditions is necessary to determine whether inefficient market rules or ISO actions may have artificially inflated prices in these hours. The following sections will focus on whether withholding by market participants or external transaction scheduling issues may have contributed to increases in these prices.

Table 4 below summarizes the supply and demand conditions in these hours, including the ISO's shortfall (reserve requirement - reserve designation) and the available undesignated resources for each class of reserve. The excess available reserves represent the total capability from each resource that can provide reserves and was not selected to provide energy or reserves in the given hour. For example, 553 MW of resources were available on July 23 at 6 p.m., of which 95 MW were peaking resources with reserve-availability bids of less than $500 per MW.

Table 4
New England Operating Reserve Conditions in High Priced Hours
Hours with ECP > $200 During Summer 2001

	Market Statistics			Reserves Short Fall			Excess Available Reserves	
Date and Time	Energy Clearing Price	Externals Accepted @ $1000	Externals Unaccepted @ $1000	10-Minute Spin	10-Minute Total	Total Reserves	All Units	Peaking Resources < $500
7/23/01 - 6 PM	$1,000	288	64	0	0	0	553	95
7/23/01 - 7 PM	$1,000	288	115	0	0	0	1,217	179
7/24/01 - 10 AM	$226	0	352	0	0	0	1,049	231
7/24/01 - 1 PM	$1,000	321	0	0	0	-53	0	13
7/24/01 - 2 PM	$1,000	334	0	0	0	-121	0	13
7/24/01 - 3 PM	$1,000	352	0	0	0	-38	0	0
7/25/01 - 12 AM	$1,000	352	0	0	0	0	199	34
7/25/01 - 1 PM	$1,000	352	0	0	0	-12	0	0
7/25/01 - 2 PM	$1,000	352	0	0	-1	-301	0	0
7/25/01 - 3 PM	$1,000	352	0	0	-6	-391	75	0
7/25/01 - 4 PM	$1,000	352	0	0	0	-115	15	0
7/25/01 - 5 PM	$1,000	352	0	0	0	-289	299	0
7/25/01 - 6 PM	$1,000	352	0	0	0	0	487	17
7/25/01 - 7 PM	$1,000	352	0	0	0	0	752	82
8/9/01 - 12 AM	$1,000	352	0	0	0	0	358	319
8/9/01 - 1 PM	$1,000	33	0	0	-9	-249	53	2
8/9/01 - 3 PM	$243	0	0	0	-548	-1,120	0	0
8/9/01 - 4 PM	$217	0	0	0	-299	-847	76	0

Source: ISO-NE Operations and Market Data. Potomac Economics analysis.

The peaking resources with availability bids of less than $500 per MW are identified because these resources are particularly suitable as reserve providers and should be economically preferred to accepting $1000 imports as a means to maintain reserves.

These data show that in most of the hours that the ISO accepted out-of-merit imports, it was unable to satisfy its reserve and energy requirements without the imports. The table shows that in only 4 hours the ISO would not have been short of reserves, even if it had not accepted the out-of-merit import transactions. In particular, the ISO accepted $1000 per MWh transactions in the hours beginning 6 p.m. and 7 p.m. on July 23 and July 25 when load was ramping down and considerable amounts of resources were available to provide reserves.

I cannot conclude from this data that the ISO erred in accepting the out-of-merit imports for two reasons. First, the transactions are scheduled 30 minutes prior to the hour and must cover the anticipated needs over the entire hour – up to 90 minutes from the time the external transaction is selected. The pricing reforms recently implemented by the ISO address this uncertainty by changing the market rules to allow the out-of-merit imports to set prices only when they are truly needed to meet either the energy or reserve obligations of the ISO.

Second, due to data limitations it is not possible to verify that all of the resources indicated as available to provide reserves could actually have been selected by the ISO. For example, some resources may have been ramp limited, subject to minimum down-time restrictions, or other restrictions that would prevent a resource from being designated as a reserve. In addition, some of these resources may have been uneconomic in comparison to the out-of-merit import transactions due to high energy or reserve bid prices, high opportunity costs, or long minimum run-times. To exclude these factors, I performed an analysis of the available resources focused exclusively on peaking generation priced competitively.

These resources were not dispatched for energy, or designated for reserves. In addition, they have start-up times of less than 30 minutes, and minimum run-times of 1 hour or less. Some of these resources qualify as TMNSR resources and could satisfy the ISO's 10-minute reserve requirements while off-line. However, all of these resources could be dispatched for energy in place of the out-of-merit imports, allowing the ISO to create additional reserves using on-line resources. Table 4 shows the results of this analysis in the final column. These results show that small quantities of these peaking resources were available to meet the energy or reserve requirements of the ISO that were not designated, but not enough to avoid accepting the out-of-merit external contracts.

Therefore, while it is possible that the scheduled out-of-merit imports were unnecessary in a limited number of hours, this report cannot conclude that the ISO erred in scheduling these imports.

B. External Transactions

The analysis in the prior section evaluated whether the out-of-merit imports were selected when economic internal resources were available. This section provides a similar analysis to assess the ISO's selection of the $1000 external imports relative to other external transactions that may have been available. The results of this analysis are shown in Table 5 below.

Table 5
New England External Transactions in High Priced Hours
Hours with ECP > $200 During Summer 2001

	Market Statistics		Import Transactions				New York Scheduling		
Date and Time	Energy Clearing Price (1)	Excess Available Reserves (2)	Net Imports (3)	Imports Accepted @ $1000 (4)	Economic Imports not Accepted from Canada (5)	Economic Imports not Accepted from NY (6)	New York Hour Ahead Price (7)	New York Real-Time Price (8)	Unaccepted Imports Economic at NY R-T (9)
7/23/01 - 6 PM	$1,000	553	2,726	288	40	0	$ 94	$ 52	0
7/23/01 - 7 PM	$1,000	1,217	2,495	288	40	0	$ 86	$ 48	0
7/24/01 - 10 AM	$226	1,049	2,330	0	0	0	$ 212	$ 81	0
7/24/01 - 1 PM	$1,000	0	3,086	321	25	0	$ 1,000	$ 69	150
7/24/01 - 2 PM	$1,000	0	3,022	334	25	0	$ 1,000	$ 67	350
7/24/01 - 3 PM	$1,000	0	3,142	352	25	0	$ 999	$ 80	400
7/25/01 - 12 AM	$1,000	199	3,298	352	144	0	$ 846	$ 262	0
7/25/01 - 1 PM	$1,000	0	3,084	352	28	0	$ 1,000	$ 187	0
7/25/01 - 2 PM	$1,000	0	2,853	352	30	0	$ 1,000	$ 664	0
7/25/01 - 3 PM	$1,000	75	2,982	352	5	0	$ 5,329	$ 364	0
7/25/01 - 4 PM	$1,000	15	3,065	352	7	0	$ 1,156	$ 321	0
7/25/01 - 5 PM	$1,000	299	2,919	352	106	0	$ 1,064	$ 174	0
7/25/01 - 6 PM	$1,000	487	2,804	352	111	0	N/A	$ 91	0
7/25/01 - 7 PM	$1,000	752	2,635	352	42	0	N/A	$ 74	0
8/9/01 - 12 AM	$1,000	358	3,087	352	48	0	$ 999	$ 255	0
8/9/01 - 1 PM	$1,000	53	2,500	33	48	0	$ 1,000	$ 233	0
8/9/01 - 3 PM	$243	0	2,034	0	48	0	N/A	$ 108	0
8/9/01 - 4 PM	$217	76	2,142	0	148	0	N/A	$ 145	0

Source: ISO-NE Transactions and Market Data, NYISO Transaction Bid Data, Potomac Economics analysis.

Column 5 of the table shows that in most of the hours there were small quantities of economic imports that were not accepted by the ISO. However, these imports were from New Brunswick and Quebec and could not be delivered to the New England load in these hours due to the internal transmission constraints. Imports from New York are not limited by these constraints and, therefore, no economic transactions should remain unaccepted in these hours from New York.

Column 6 of the table confirms that ISO New England did, in fact, schedule all economic transactions that were available to it. However, this does not mean that all economic transactions were scheduled between New York and New England because the New York transaction scheduling process did not allow all economic transactions to be scheduled.

Columns 7 and 8 show the New York prices at the New England border in the real-time market versus the prices produced by New York's hour ahead model – the balancing market evaluation or BME. All transactions are scheduled hour-ahead by New York and New England. New York's BME schedules transactions on the basis of the bids it receives with the transactions. An export from New York to New England is treated as a load bid at the border. Therefore, the transaction will be scheduled in any hour that the BME price is less than the bid of the exporting participant. In the peak hours examined in this section, the hour-ahead prices produced by the BME did not serve as a reliable forecast for the real-time price. This caused lower cost exports from New York to fail to be scheduled by the New York ISO in some hours.

Column 9 shows the quantities that would have been scheduled by the New York ISO if the real-time price had prevailed in the BME. This issue contributed in these hours to the ISO New England's need to accept the $1000 imports. The New York ISO has been working with stakeholders to identify the factors that have caused the BME prices and real-time prices to diverge in peak hours. Several causes have been identified and should be remedied prior to this summer. These remedies will allow increased exports from New York to New England when New England is in shortage conditions (unless New York is experiencing comparable shortages).

C. Potential Withholding in High-Priced Hours

The final assessment evaluates the extent to which the conduct of participants may have contributed to the high prices in these hours. This contribution would come in one of three forms: physical withholding of capacity, economic withholding of energy, or economic withholding of reserve capability.

To evaluate the first two forms of withholding, the report utilizes the statistics that are analyzed for the entire year in the prior sections, namely the output gap for economic withholding and various forms of outages and deratings for physical withholding. The levels of these statistics in the high-priced hours are shown in Table 6 below.

Table 6

Potential Withholding in High Priced Hours

Hours with ECP > $200 During Summer 2001

	Market Statistics		Outages and Deratings (%)				Reserve Bids > $900/MW	
Date and Time	**Energy Clearing Price**	**Externals Accepted @ $1000**	**Multiday Forced Outages**	**Intraday Forced Outages**	**Other Deratings**	**Output Gap**	**TMNSR**	**TMOR**
7/23/01 - 6 PM	$1,000	288	3%	0%	11%	0.3%	17	17
7/23/01 - 7 PM	$1,000	288	3%	0%	11%	0.4%	17	17
7/24/01 - 10 AM	$226	-	5%	0%	7%	0.1%		
7/24/01 - 1 PM	$1,000	321	5%	0%	6%	0.0%	47	112
7/24/01 - 2 PM	$1,000	334	5%	0%	6%	0.0%	28	128
7/24/01 - 3 PM	$1,000	352	5%	0%	6%	0.0%	18	141
7/25/01 - 12 AM	$1,000	352	3%	0%	6%	0.0%	144	27
7/25/01 - 1 PM	$1,000	352	3%	0%	6%	0.0%	50	121
7/25/01 - 2 PM	$1,000	352	2%	0%	6%	0.0%	17	79
7/25/01 - 3 PM	$1,000	352	2%	0%	5%	0.0%	0	87
7/25/01 - 4 PM	$1,000	352	2%	0%	6%	0.0%	11	83
7/25/01 - 5 PM	$1,000	352	3%	0%	7%	0.0%	11	81
7/25/01 - 6 PM	$1,000	352	3%	0%	7%	0.3%	28	90
7/25/01 - 7 PM	$1,000	352	4%	0%	8%	0.9%	28	132
8/9/01 - 12 AM	$1,000	352	2%	0%	4%	0.0%	20	17
8/9/01 - 1 PM	$1,000	33	2%	0%	5%	0.0%	1	22
8/9/01 - 3 PM	$243	-	2%	0%	6%	0.1%		
8/9/01 - 4 PM	$217	-	2%	0%	6%	0.1%		

Source: ISO-NE Operations and Market Data. Potomac Economics analysis.

This table shows that the outages and deratings, in total, ranged from 6 percent to 14 percent of the market capacity. The average in all hours for these values is approximately 8 percent. 11 of the 18 high-priced hours exhibit deratings higher than this level while 7 hours are at or below this level. While these results suggest that some degree of physical withholding could have occurred in these hours and contributed to the tight market conditions in these hours, it is impossible to draw this conclusion from these isolated hours.

The primary means to evaluate discrete instances of potential physical withholding is through an audit program as part of the market monitoring effort. The statistical analysis in this report may assist in focusing the audits on the most likely instances, but the audits themselves are necessary to conclude that an outage has been claimed for strategic purposes.

The output gap results shown in this table indicate that economic withholding did not play an important role setting these prices. For purposes of this table, I computed the output gap assuming that the ECP was $950 to ensure that any internal units raising their energy bid to close to $1000 per MWh and contributing to the ISO's decision to accept $1000 imports would be identified. [27]

The only potential significant quantities shown in the output gap occur in the 6 p.m. and 7 p.m. hours on July 23 and July 25. Other deratings also rise moderately in these hours. An examination of the data in these hours indicates that these quantities are primarily associated with pumped storage and reservoir hydro units. By the 6 p.m. hour, it is not unreasonable to expect that the hydro resources will begin to curtail their output to preserve the ability to produce in future periods when the power is even more valuable. This is sometimes accomplished by raising the bid price for the unit, which will cause it to appear in the output gap.

These output considerations are particularly relevant for pumped storage that must stop producing when its reservoir decreases to a particular level. When this occurs, the unit will be derated to reflect that it must pump to refill its reservoir before it produces additional output.

The final results shown in the table are the quantity of TMNSR and TMOR with bids above $900 per MWh, including both accepted and unaccepted quantities. Bids at these levels may be considered economic withholding to the extent that the marginal costs of providing the reserves are substantially lower. This should be the case since the reserve designations are made *ex post* -- it is unlikely that the units incur significant marginal costs associated with providing reserves.

However, even if these bids do represent economic withholding, their quantities are substantially less in all hours than the quantity of $1000 imports accepted by the ISO. Therefore, one may conclude that $1000 per MWh prices, but for these reserve bids, would have remained at $1000 per MWh.

D. Conclusions Regarding the High-Priced Hours During 2001

A number of conclusions can be drawn from the analysis presented in this section. First, in the majority of the high-priced hours, the ISO was sufficiently deficient of internal resources that accepting the $1000 per MWh imports was warranted. However, this deficiency may not have prevailed for the entire hour so the $1000 per MWh price may have been justified for only a portion of some of these hours. The recently implemented pricing reforms will address this issue, as well as cases where the ISO accepts an out-of-merit import uneconomically due to uncertainty regarding its need for the import at the time that it is accepted.

In addition, it is not clear that the economic value of reserves, and reliability more broadly, has been established and reflected appropriately in the current New England markets. If the reserves were not worth $1000 per MW that the ISO sought to maintain, the acceptance of the imports and associated energy price may not have been economically justified. An infinite value for reserves is implicit in the current market rules, given the fact that the reserve requirement is an absolute requirement. However, NEPOOL may consider over the longer-term implementing a demand curve for reserves, establishing an explicit value for reserves that would govern the ISO's actions to maintain the reserves and the resulting energy prices. Such a demand curve would reflect the increasing marginal value of reserves as the quantity of reserves falls (and the deficiency grows).

Second, the New York ISO market rules related to scheduled exports from New York to New England did restrict New England's access to lower cost imports in a few hours. If unaddressed, these conditions would likely reoccur under peak-demand conditions this summer. However, the New York ISO is implementing changes to its market models that should minimize this possibility in the future. This issue should continue to be monitored to ensure that these changes have been effective.

Lastly, no clear evidence of economic or physical withholding during these high-priced hours emerged from this analysis. However, it is important to continue to monitor for these issues, particularly in the peak-demand hours when the presence of market power is most likely. This monitoring should include the types of screening and analysis of withholding presented in this report and, in the case of physical withholding, should be complemented by random physical audits to verify the technical justifications accompanying forced outages and significant deratings.

Appendix A

Supply Function Equilibrium

Klemperer and Meyer's Supply Function Equilibrium results described in Section II of this report are derived in this Appendix.

Assume the market contains an arbitrary number of firms. Each firm chooses a supply function indicating how much it is willing to supply at each price level, p. Denote the supply function of a representative firm (firm i), as $S_i(p)$. Each firm experiences increasing costs according to a cost function C: at price p, firm i produces $S_i(p)$ so that it has total costs denoted as $C_i[S_i(p)]$.

If market demand is given by D, then firm i faces residual demand $D – S_{-i}(p)$, where $S_{-i}(p)$ denotes the sum of all other firms (except for firm i) supply-function output at price p.. Hence, firm i's profit function is given by:

(A1) $$\pi(p)_i = p * [(D - S_{-i}(p)] – C_i[D – S_{-i}(p)].$$

To achieve an optimal supply function in light of (A1), firm i maximizes (A1) with respect to p:

(A2) $$0 = D - S_{-i}(p) + \{p - \partial C_i[D – S_{-i}(p))/\partial (S_{-i}(p)]\}*[\partial D/\partial p - \partial S_{-i}(p)/\partial p].$$

Since $\partial C_i[D – S_{-i}(p))/\partial (S_{-i}(p)]$ is simply the marginal cost when firm i produces $D – S_{-i}(p)$ units, we can write (A2) as:

(A3) $$0 = D - S_{-i}(p) + \{p - MC[D - S_{-i}(p)]\}*[\partial D/\partial p - \partial S_{-i}(p)/\partial p].$$

In equilibrium, market supply equals market demand. Since what all other firms produce (except firm i) is $S_{-i}(p)$, and firm i produces $S_i(p)$, market supply = $S_{-i}(p)+ S_i(p)$. Hence, the equilibrium requires $D = S_i(p) + S_{-i}(p)$ which implies $D - S_{-i}(p) = S_i(p)$ and (A3) becomes:

(A4) $$0 = S_i(p) + \{p - MC[S_i(p)]\}*[\partial D/\partial p - \partial S_{-i}(p)/\partial p].$$

Assuming a symmetric equilibrium (i.e., $S_i(p) = S_j(p) = S(p)$ for all firms i, j (which Klemperer and Meyer prove exists), then (A4) becomes:

(A5) $$0 = S(p) + \{p - MC[S(p)]\}*[\partial D/\partial p - \partial S(p)/\partial p].$$

Let q* denote the equilibrium supply, $S(p)$. Also, because electricity demand is highly price inelastic, let dD/dp = 0. Using the inverse of the supply function (and invoking the Implicit Function Theorem), the optimal level of output for each firm is expressed in the following equation, which represents the first order condition for each supplier:

(A6) $$q^* = (p – MC(q^*)) / \partial p/\partial S.$$

Under perfect competition, the firm will be a price taker, producing until the price equals its marginal cost. Let q^c be the competitive level of output that solves p = MC(q).

Hence, $q^c - q^*$ would be the divergence of the firm's optimal output from the competitive output level, defined as its optimal level of withholding. Therefore, for all q^* that solve (A6), the optimal level of withholding can be expressed as:

(A7) $$w^* = q^c - (p - MC(q^*)) / \partial p / \partial S.$$

This result shown in (A7) is the basis for the empirical hypotheses described in Section II of the report that serve as the basis of the withholding analysis.

Appendix B

Estimation of Reference Prices

The competitive benchmark used in the analysis is a critical element in estimating the output gap because it determines the economic level of output for a unit at a given ECP. This section describes how this benchmark is developed for the analysis of economic withholding in this report.

In a competitive single-price auction where no supplier can influence the price, suppliers maximize their profits by accepting the clearing price when that price is higher than their marginal costs of producing and by not producing when that price is lower (i.e., behaving as a price taker). The competitive conduct implied by this assertion is that when a supplier lacks market power it should be bidding its marginal costs into the auction market. Therefore, the competitive benchmark can be estimated by using the historical accepted bids from the resource to serve as a proxy for each resource's marginal costs.

The primary competitive benchmark used in this report is a reference price that is generally calculated based on an average of bids accepted in-merit for each unit. If, as described in the introduction, withholding is not rational under most market conditions, then taking an average from a supplier's accepted in-merit bids during a long time period should provide a reliable indicator of the competitive bid level for the unit. Reference values are calculated for the entire output range of the unit (in 1, 2, 5, or 10 MW segments depending on the size of the unit). The referenced value is the lower of the mean or median of the accepted in-merit bids for each respective output segment with an adjustment described below for fuel price changes. Accepted bids are said to be in-merit at output levels where the incremental bid price is at or below the ECP. The reference price is calculated using a 120-day season with Summer defined as June to September.

An adjustment is made to the average accepted bid for each unit based on spot fuel prices two days before each bid was accepted – this is the fuel price seen by the participant when it develops its bid. The fuel price adjustment is made by changing 90% of the average reference price over the season by the ratio of the current fuel price to the seasonal average fuel price. In other words, it is assumed that 90% of the price is attributable to fuel costs. Hence this 90% segment is adjusted for the difference between the current fuel price and the average over the season.

No-load bids and self-scheduled bids must be treated differently in the calculation of the competitive benchmark. The combination of a no-load bid with an incremental bid curve implies a minimum average running cost per MWh for the unit. When individual bid curve segments fall below this implied minimum, this minimum is substituted for those segments in the calculation of accepted bids.

A self schedule is essentially a bid to produce energy regardless of the ECP. Therefore the bid curve segments at output levels below the self-scheduled quantity provide little insight into the marginal costs for that unit, and they are not used in the competitive benchmark calculation. Bids of $1.00 or less are effectively the same as self schedules and do not give us any information about the production costs of the unit; thus they are excluded from the calculations. In many cases, units submit no-load bids and self-scheduled bids simultaneously. Since the ISO does not compensate self-scheduled units according to their no-load bids, these units are also excluded from the calculation of the competitive benchmark.

The primary competitive benchmark may be inadequate for certain ranges of a particular unit when it is based on a very small number of accepted in-merit bids. This may occur because the output segment was out-of-merit very frequently, did not run very many times, or was usually self-scheduled. This primary benchmark is considered inadequate for an individual output segment if it is based on fewer than ten accepted bids. When there are too few accepted bids for a segment, two alternative measurements are used to calculate a secondary reference price to be used in the analysis as the competitive benchmark.

The first alternative reference price is used for self-scheduled resources, which we define as including resources bidding $1.00 per MWh or less. Participants bidding in this manner when they are reasonably confident that the ECP will exceed their marginal costs are likely to not be providing reliable information regarding their marginal costs. Therefore, the alternative reference price for these resources is computed as the lower of the mean or median of the lowest quartile of ECPs in hours where the output segment was self-scheduled. A fuel price adjustment similar to the one used for the primary competitive benchmark is made before selecting the hours that fall into the lowest quartile.

The second alternative to the primary competitive benchmark is based on all bids provided by the unit, whether accepted or not. Again, the lower of the mean or median of the lowest quartile of fuel price-adjusted bids is used to calculate a reference price for each output segment. Thus, if a unit typically bids in excess of its marginal costs, it would have to do so in the vast majority of the hours (in excess of 75%) for this average to rise significantly above its marginal costs. For this alternative competitive benchmark, no-load bids and self-scheduled bids are taken into account in the same manner as in determining the primary competitive benchmark.

These two alternative measures are combined into a composite reference price for each output segment when the primary competitive benchmark is not available. For each output segment, the two alternatives are combined into an average that is weighted by the number of hours in which each alternative is used. Thus, if a unit is self-scheduled at a particular output level in 30% of the hours in the season, the composite reference price would be the average of the first and second alternative measures weighted by 30 percent and 70 percent, respectively.

The primary competitive benchmark is used for output segments where it is available. When it is not available, the secondary competitive benchmark is used. Reference price segments at each output level are combined into a single step curve, which is constrained to be monotonically increasing. If the segments are not monotonically increasing, an algorithm is used that redistributes the local peaks while holding constant the total area under the reference price curve.

Appendix C

Linear Regression Procedure and Results

The linear regression model is intended to determine the extent to which various measures of potential withholding (e.g., output gap, deratings) can be explained by a combination of market factors and participant-specific factors, generally referred to as "explanatory" variables or independent variables. These factors include both strategic and non-strategic factors as described in the report. The regression procedure estimates a coefficient for each explanatory variable and calculates the statistical properties of the coefficient that can be used to test the statistical significance of the particular variable in relation to the dependent variable. The estimated coefficient can be interpreted as the incremental change in the withholding measure resulting from a one unit change in the explanatory variable.

Our basic regression model can be expressed as

(C1) $Y_t = a_0 + a_1{*}Y_{t-1} + b_1{*}X_1 + b_2{*}X_2 + \ldots + b_n{*}X_n + d_1Z_1 + d_2Z_2 + \ldots + d_mZ_m + e$

where:

- Y_t is the given measure of economic or physical withholding in period t
- a_0 is a constant.
- Y_{t-1} is the dependent variable lagged one period with a coefficient of a_1.
- X_1 through X_n are the n variables that are hypothesized to affect the non-strategic level of Y_t.
- b_1 through b_n are the associated coefficients of the X_1 through X_n explanatory variables. Hence, b_1 indicates the amount by which Y_t increases when X_1 increases, etc.
- Z_1 through Z_m are the m variables that are hypothesized to affect the strategic level of Y_t. In this model, these variables relate to the demand level and size of participant.
- d_1 through d_m are the estimated coefficients of the Z_1 through Z_m.

- e is the disturbance term in the equation and contains the effects of unobservable phenomena. Importantly, on average, e is assumed to be zero.[28]

Given the presence of the disturbance term, the estimated coefficient will be equal to the true coefficient only probabilistically. Indeed, the estimated coefficient is distributed around a mean that is equal to the true value of the coefficient. The variability of the estimate around the mean is measured by the standard error. This is calculated based on the variability of the explanatory variable itself, other explanatory variables, and the dependent variable. The ratio of the standard error and the estimated coefficient is distributed in accordance with the *Student's t-distribution*[29] and is called the *t-statistic* or *t* value. Therefore, the *t*-statistic allows the investigator to determine whether the estimated coefficient is statistically significant at a given confidence level and, thus, whether it possesses a statistically meaningful relationship to the dependent variable.

A. Output Gap

The selection of the explanatory variables, is called model specification, is an important aspect of the regression procedure. The key element of model specification is to identify variables that constitute a viable economic model that would explain the dependent variable and satisfy the requirements of classical OLS.[30]

The dependent variable in the first analysis is the output gap for each participant. This value is normalized for differences in participants' sizes by dividing the gap by the participant's total capacity. This normalized value, as opposed to the absolute output gap level for each participant, is used to satisfy the classical assumptions of OLS. Under OLS, the variance of disturbance should not vary across the cross section, but it would vary if one did not account for the widely varying sizes of the participants.

As described in the body of the report, two types of explanatory variables are used to evaluate whether the output gap represents an attempt economically to withhold capacity from the market. The first set contains variables that are expected to influence the level of the output gap in a non-strategic manner. The following is a description of the non-strategic variables.

AGE	The average age of the plants in the market participant's portfolio weighted by plant size.

FOSSIL SHARE	The share of base-load fossil units in the market participant's portfolio.
PEAKER SHARE	The share of fossil peaking units in the market participant's portfolio, such as gas turbines.
HYDRO SHARE	The share of hydro plants in a market participant's portfolio.
FO6 PRICE	The price of Fuel Oil number 6 in New England.
NG PRICE	The price of natural gas in New England.
SUMMER	A dummy variable indicating whether the hour occurs during the summer.
WINTER	A dummy variable indicating whether the hour occurs during the winter.
WORK HOURS	A qualitative variable indicating whether or not the observation is in a hour from 6AM to 10PM on weekdays that are non-holidays.
OOM SHARE	The proportion of the participant's portfolio that is dispatched out-of-merit more than 20% of the time when it is dispatched.

The second category of explanatory variables contains the variables that identify factors that may be related to the presence of or incentive to exercise market power, generally referred to in this report as strategic variables. These variables include:

LARGE PARTICIPANT	Large Market Participant -- a qualitative variable equal to 1 when a participant's portfolio is greater than 1200 MW.
PEAK DEMAND	A qualitative variable equal to 1 in an hour when the demand in the hour is in the highest 1 percent of demand levels.
LARGE@PEAK	A qualitative variable that is equal to 1 when both *LARGE PARTICIPANT* and *PEAK DEMAND* equal 1. This will indicate the combined conditions of large participant and high demand.

The results of the estimated model are presented in Table C1. This table shows the coefficients estimated for each variable and the *p*-value for each estimate. The asterisk next to a *p*-value indicates that the estimate is statistically significant at the 95 percent confidence level (i.e., the *p*-value is less than 0.05).

Table C1 -- Regression Results: Economic Withholding
Dependent Variable: Output Gap Share of Portfolio

Variable	Base Model Estimate	Base Model *p-value*	3 PM Hour Only Estimate	3 PM Hour Only *p-value*	Fossil Units Estimate	Fossil Units *p-value*
Non-Strategic Variables						
Intercept	-0.0021	0.007 *	-0.0042	0.373	-0.0018	0.154
LAGGED DEPENDENT	0.7035	0.000 *	0.4934	0.000 *	0.4443	0.000 *
AGE	-0.0002	0.000 *	-0.0004	0.000 *	-0.0002	0.000 *
FOSSIL STEAM SHARE	0.0075	0.000 *	0.0112	0.000 *	0.0083	0.000 *
PEAKER SHARE	-0.0008	0.051	-0.0030	0.194		
HYDRO SHARE	0.0054	0.000 *	0.0130	0.001 *		
FO6PRICE	-0.0005	0.064	-0.0004	0.802	-0.0011	0.003 *
NGPRICE	0.0009	0.000 *	0.0020	0.000 *	0.0012	0.000 *
SUMMER	0.0014	0.000 *	0.0061	0.000 *	0.0017	0.000 *
WINTER	0.0042	0.000 *	0.0053	0.002 *	0.0056	0.000 *
WORK HOURS	0.0007	0.002 *	-0.0031	0.023 *	-0.0001	0.864
OOM SHARE	0.0110	0.000 *	0.0239	0.000 *	0.0103	0.000 *
Strategic Variables						
LARGE PARTICIPANT	-0.0002	0.471	0.0020	0.276	-0.0015	0.001 *
PEAK DEMAND	-0.0031	0.008 *	-0.0138	0.002 *	-0.0047	0.010 *
LARGE@PEAK	0.0003	0.913	0.0030	0.784	0.0031	0.462
Adjusted R^2	0.53		0.30		0.49	
Observations	339,844		14,057		237,342	

Notes: Base model estimated using all hours in 2001 and economic withholding using 110% of reference prices in the withholding screen. The Fossil Units case uses only fossil units. To address autoregression in the disturbance term, the one-period lag of the dependent variable was included in the model and was statistically significant.
* Statistically significant at the .05 level.

The table shows three cases: 1) a case including all hours and all units, 2) a case including only the 3 p.m. hour in each day, and 3) a case including all hours, but only the fossil-fired generation owned by each participant.

B. Physical Withholding

To test for strategic physical withholding, we estimated a model analogous to the model used to test for strategic economic withholding with a few changes to some of the explanatory variables. In particular, we omitted fuel prices and out-of-merit shares because these variables provide no intuitive explanation for either forced outages or other deratings.

While the economic withholding analysis includes a number of cases with a single dependent variable (the normalized output gap) to evaluate physical withholding we analyze the determinant of four different dependent variables. The first three models are shown in Table C2.

Table C2 -- Regression Results: Physical Withholding
Outages and Derating Share of Portfolio - 3 p.m. Hours in 2001

	Base Model		Other Deratings		Forced Outages	
Variable	**Estimate**	***p-value***	**Estimate**	***p-value***	**Estimate**	***p-value***
Non-Strategic Variables						
Intercept	0.0409	0.000 *	0.0382	0.000 *	0.0041	0.005 *
LAGGED DEPENDENT	0.6958	0.000 *	0.6918	0.000 *	0.6087	0.000 *
AGE	-0.0001	0.156	-0.0002	0.020 *	0.0001	0.110
FOSSIL STEAM SHARE	-0.0072	0.007 *	-0.0060	0.014 *	-0.0015	0.282
PEAKER SHARE	-0.0026	0.447	0.0008	0.808	-0.0041	0.018 *
HYDRO SHARE	0.0745	0.000 *	0.0841	0.000 *	-0.0107	0.001 *
SUMMER	-0.0048	0.041 *	-0.0051	0.018 *	0.0003	0.832
WINTER	-0.0071	0.005 *	-0.0072	0.002 *	0.0000	0.972
WORK HOURS	-0.0101	0.000 *	-0.0107	0.000 *	0.0006	0.613
Strategic Variables						
LARGE PARTICIPANT	-0.0202	0.000 *	-0.0232	0.000 *	0.0033	0.029 *
PEAK DEMAND	-0.0030	0.669	-0.0093	0.146	0.0062	0.090
LARGE@PEAK	0.0022	0.899	0.0091	0.568	-0.0060	0.504
Adjusted R^2	0.55		0.56		0.37	
Observations	14,057		14,057		14,057	

Notes: Total deratings is the dependent variable in the Base Model. For the Other Deratings case, the dependent variable is total deratings excluding forced outages. For the Forced Outage case, the dependent variable is forced outages of up to 7 days in length. To address autoregression in the error term, a one-period lag of the dependent variable was included and was statistically significant in all cases.

* Statistically significant at the .05 level.

The first model, the "Base Model", uses total deratings as the dependent variable. "Total deratings" includes all differences between each unit's maximum capability and its upper operating limit in an hour, excluding planned and long-term forced outages. The second model includes only "other deratings" in the dependent variable, which is computed by subtracting the short-term forced outages from the total deratings. The third model uses only the short-term forced outages as the dependent variable.

The last model includes in the dependent variable only the intraday forced outages, which are a subset of the short-term forced outages. These results are shown in Table C3 together with the Base Model.

Table C3 -- Regression Results: Intraday Forced Outages
Outages and Derating Share of Portfolio - 3 p.m. Hours in 2001

Variable	Base Model Estimate	Base Model *p-value*	Intraday Forced Outages Estimate	Intraday Forced Outages *p-value*
Non-Strategic Variables				
Intercept	0.0409	0.000 *	-0.0004	0.234
LAGGED DEPENDENT	0.6958	0.000 *	0.0506	0.000 *
AGE	-0.0001	0.156	0.0000	0.052
FOSSIL STEAM SHARE	-0.0072	0.007 *	0.0003	0.255
PEAKER SHARE	-0.0026	0.447	-0.0005	0.226
HYDRO SHARE	0.0745	0.000 *	-0.0010	0.132
SUMMER	-0.0048	0.041 *	0.0004	0.190
WINTER	-0.0071	0.005 *	-0.0002	0.519
WORK HOURS	-0.0101	0.000 *	0.0004	0.109
Strategic Variables				
LARGE PARTICIPANT	-0.0202	0.000	-0.00004	0.905
PEAK DEMAND	-0.0030	0.669	0.0030	0.000 *
LARGE@PEAK	0.0022	0.899	-0.0033	0.093
Adjusted R^2	0.55		0.004	
Observations	14,057		14,057	

Notes : Total deratings is the dependent variable in the Base Model. For the Other Deratings case, the dependent variable is total deratings excluding forced outages. For the Forced Outage case, the dependent variable is forced outages of up to 7 days in length. To address autoregression in the error term, a one-period lag of the dependent variable was included and was statistically significant in all cases.

* Statistically significant at the .05 level.

Endnotes

1 Patton, D. (November 2001), "An Assessment of Peak Energy Pricing in New England During Summer 2001," Report to ISO New England.

2 Response of ISO New England Inc., F.E.R.C. Docket No. ER02-1149-000

3 Bushnell, J. and C. Saravia (February 2002), "An Empirical Assessment of the Competitiveness of the New England Electricity Market," Report to ISO-New England.

4 Dispatch is established at least every five minutes, but it could be more frequent when conditions dictate, e.g., during times of rapidly changing demand. When the dispatch is every five minutes, each RTMP is weighted equally at 1/12. When the dispatch is more frequent, the weights are in accordance with the duration of the RTMP.

5 As noted above, New England is in the process of reforming certain market rules, including introducing a multi-settlement system whereby participants can commit to day-ahead prices for any or for all hours and for all or a portion of their load or resources for the next day. This is positive because it will provide a more complete ability to hedge against risk.

6 See "Working Paper on Standardized Transmission Service and Wholesale Electric Market Design", Docket No. RM01-12-000 (March 2002).

7 Markets generally clear on a five minute to one hour basis. To ensure that supply equals demand in shorter timeframes, control areas generally send automated signals to generators equipped with Automated Generator Control equipment ("AGC").

8 Industrial organization economists will recognize physical withholding as an element of *Cournot* competition where each competitor strategically determines its quantity given the expected reaction of the other competitors. As explained more below, the concept of supply-function equilibria pioneered by Klemperer and Meyer provides a useful way to examine market dynamics when both quantity (physical withholding) and price (economic withholding) are strategically employed in the supply-side of a market.

9 See, *e.g.*, Borenstein, S., J. Bushnell, and F. Wolak (August 2000), "Diagnosing Market Power in California's Restructured Wholesale Electricity Market," University of California Energy Institute Working Paper PWP-064; Joskow, P., and E. Kahn (March 2001), "A Quantitative Analysis of Pricing Behavior In California's Wholesale Electricity Market During Summer 2000," NBER Working Paper 8157; Harvey, S. and W. Hogan (April 2001), "On the Exercise of Market Power Through Strategic Withholding in California, mimeo; P. Joskow and E. Kahn (July 2001), "Identifying the Exercise of Market Power: Refining the Estimates," mimeo; Joskow, P. and E. Kahn (February 2002), "Identifying the Exercise of Market Power: Refining the Estimates – The Final Word," mimeo.

10 See, *e.g.*, Borenstein, S., *et al*., *op. cit.*; Joskow, P. and E. Kahn (February 2002) *op. cit.*; and Bushnell, J. and C. Saravia , *op. cit*.

11 See Patton, D. (April 2001) "Annual Assessment of the New York Electric Markets", Report to the Board of Directors and Management Committee.; Patton, D. (2001) "Annual Report of the New York Independent System Operator"; Report to New York ISO Board of Directors and Management Committee; Joskow, P. and E. Kahn, *op. cit.*

12 Klemperer, P. and M. Meyer (November 1989), "Supply Function Equilibria in Oligopoly under Uncertainty," *Econometrica* 57, 1243-1277.

13 Note that one such supply function is to offer each block of capacity at marginal cost. If all suppliers offer such a supply function, a perfectly competitive outcome is obtained. Whether such a strategy by all players is an equilibrium (i.e., whether no supplier can do better by deviating from such a strategy) will depend on market structure parameters.

14 See, Bohn, R., A. Klevorick, and C. Stalon (1999), "Second Report on Market Issues in the California Power Exchange Energy Markets"; Green, R. (1996), "Increasing Competition in the British Electricity Spot Market," *The Journal of Industrial Economics*, 44, 205-216; Green, R. and D. Newberry (1992), "Competition in the British Electricity Spot Market", *Journal of Political Economy*, 100, 929-953.

15 In the Klemperer and Meyer formulation, all firms have an identical portfolio of capacity. Hence when all firms satisfy (1), $C'(q)$ is the marginal cost of the highest-cost unit for each supplier which, given identical firms, is the marginal cost of the most expensive unit dispatched in the market.

16 The basic insights provided by the supply function equilibria have been produced by others using slightly different concepts. See, Bohn, R., *et al., op. cit.*; Joskow, P. and E. Kahn, *op. cit.*

17 This is true because q^C satisfies $p = MC(q^C)$. Making the reasonable assumption that MC is higher at higher levels of output, any higher level of output beyond q^C would result in $P < MC$, something that would cause short-run operating losses. No firm would offer a supply function that permits its units to run at a price lower than MC.

18 Joskow & Kahn (2002), *op. cit.*

19 One reason for this is the higher operating pressures that increase the probability of experiencing tube leaks and other technical problems.

20 Kahn, A, P. Cramton, R. Porter, and R. Tabors (January 2001), "Blue Ribbon Panel Report: Pricing in the California Power Exchange Electricity Market: Should California Switch from Uniform Pricing to Pay-as-Bid Pricing?"

21 See Appendix C for a discussion of the calculation of confidence levels.

22 As indicated in Appendix C, the analysis implicitly controls for the share of nuclear capacity, too.

23 Since the coefficient of ***LARGE@PEAK*** is not statistically significant, the regression analysis does not provide any evidence that the conduct of large participants is different in the highest demand periods.

24 Joskow, P. and E. Kahn (2001), *op. cit.*

25 The AGE variable is significant at a 94 percent confidence level, indicating that older portfolios experience a higher share of intra-day forced outages.

26 Dispatchable external transactions are those that are scheduled by the ISO on the basis of their bid price. These transactions are scheduled on an hourly basis rather than the 5-minute timeframe on which most internal resources are dispatched.

27 The $950 per MWh price was selected as an arbitrarily high value to detect economic withholding by suppliers bidding at or very close to $1000. I ran the same test with values ranging from $900 to $975 and obtained identical results.

28 Technically, the disturbance or error term is assumed to be distributed in accordance with the Normal Distribution with a mean of zero.

29 The *t* distributions were discovered by William S. Gosset in 1908. Gosset was a statistician employed by the Guinness brewing company which had stipulated that he not publish under his own name. He therefore wrote under the pen name ``Student."

30 The classical assumptions are (1) the disturbance term is normally distributed; (2) the disturbance term has a zero mean; (3) the variance of the disturbance term is the same for all observations; (4) The disturbance term is uncorrelated between observations; and (5) the explanatory variables are nonstochastic.